

Filed: January 11, 2002 Atty. Docket No.: 015389-002640US  
Applicant: Thomas R. Cech et al.  
Title: HUMAN TELOMERASE CATALYTIC SUBUNIT:  
DIAGNOSTIC AND THERAPEUTIC METHODS  
**COPY OF SEQUENCE LISTING**  
**FROM PARENT APPLICATION NO. 08/912,951**

(1) GENERAL INFORMATION:

- (i) APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.
- (ii) TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS
- (iii) NUMBER OF SEQUENCES: 335
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESS: Townsend and Townsend and Crew LLP
  - (B) STREET: Two Embarcadero Center, 8th Floor
  - (C) CITY: San Francisco
  - (D) STATE: California
  - (E) COUNTRY: United States of America
  - (F) ZIP: 94111
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/912,951
  - (B) FILING DATE: 14-AUG-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/854,050
  - (B) FILING DATE: 09-MAY-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/851,843
  - (B) FILING DATE: 06-MAY-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/846,017
  - (B) FILING DATE: 25-APR-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/844,419
  - (B) FILING DATE: 18-APR-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/724,643
  - (B) FILING DATE: 01-OCT-1996
  - (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Apple, Randolph T.
- (B) REGISTRATION NUMBER: 36,429
- (C) REFERENCE/DOCKET NUMBER: 015389-002600US

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (415) 576-0200
- (B) TELEFAX: (415) 576-0300

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4015 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 56..3454
- (D) OTHER INFORMATION: /product= "hTERT"  
/note= "human telomerase reverse  
transcriptase (hTERT) catalytic protein  
component"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC CCGCG ATG 58  
Met  
1

CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC CAC 106  
Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser His  
5 10 15

TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG GGG CCC 154  
Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly Pro  
20 25 30

CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT TTC CGC GCG 202  
Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg Ala  
35 40 45

CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC GCA CGG CCG CCC 250  
Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro Pro  
50 55 60 65

CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC CTG AAG GAG CTG GTG 298  
Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu Val  
70 75 80

GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC GGC GCG AAG AAC GTG CTG 346  
Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val Leu  
85 90 95

GCC TTC GGC TTC GCG CTG CTG GAC GGG GCC CGC GGG GGC CCC CCC GAG 394  
Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Gly Pro Pro Glu  
100 105 110

GCC TTC ACC ACC AGC GTG CGC AGC TAC CTG CCC AAC ACG GTG ACC GAC Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr Asp 115 120 125	442
GCA CTG CCG GGG AGC GGG GCG TGG GGG CTG CTG CTG CGC CGC GTG GGC Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val Gly 130 135 140 145	490
GAC GAC GTG CTG GTT CAC CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val Leu 150 155 160	538
GTG GCT CCC AGC TGC GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr Gln 165 170 175	586
CTC GGC GCT GCC ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro His Ala Ser Gly Pro 180 185 190	634
CGA AGG CGT CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg Glu 195 200 205	682
GCC GGG GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg Gly 210 215 220 225	730
GGC AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg Gly 230 235 240	778
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG GCC Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp Ala 245 250 255	826
CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT GTG GTG His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val Val 260 265 270	874
TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG GGT GCG CTC Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala Leu 275 280 285	922
TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC CAG CAC CAC GCG Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His Ala 290 295 300 305	970
GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC TGG GAC ACG CCT TGT Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro Cys 310 315 320	1018
CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC CTC TAC TCC TCA GGC GAC Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly Asp 325 330 335	1066
AAG GAG CAG CTG CGG CCC TCC TTC CTA CTC AGC TCT CTG AGG CCC AGC Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro Ser 340 345 350	1114
CTG ACT GGC GCT CGG AGG CTC GTG GAG ACC ATC TTT CTG GGT TCC AGG Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser Arg 355	1162

355	360	365	
CCC TGG ATG CCA GGG ACT CCC CGC AGG TTG CCC CGC CTG CCC CAG CGC Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln Arg 370 375 380 385			1210
TAC TGG CAA ATG CGG CCC CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His Ala 390 395 400			1258
CAG TGC CCC TAC GGG GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg Ala 405 410 415			1306
GCG GTC ACC CCA GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln Gly 420 425 430			1354
TCT GTG GCG GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp Pro Arg Arg Leu Val 435 440 445			1402
CAG CTG CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe Val 450 455 460 465			1450
CGG GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser Arg 470 475 480			1498
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC CTG His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser Leu 485 490 495			1546
GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG ATG AGC Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met Ser 500 505 510			1594
GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT GGC TGT GTT Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys Val 515 520 525			1642
CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG GCC AAG TTC CTG Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe Leu 530 535 540 545			1690
CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG CTC AGG TCT TTC TTT His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe 550 555 560			1738
TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG CTC TTT TTC TAC CGG Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg 565 570 575			1786
AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT GGA ATC AGA CAG CAC TTG Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu 580 585 590			1834
AAG AGG GTG CAG CTG CGG GAG CTG TCG GAA GCA GAG GTC AGG CAG CAT Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln His 595 600 605			1882



CGG GAA GCC AGG CCC GCC CTG CTG ACG TCC AGA CTC CGC TTC ATC CCC Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro 610 615 620 625	1930
AAG CCT GAC GGG CTG CGG CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val Gly 630 635 640	1978
GCC AGA ACG TTC CGC AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg 645 650 655	2026
GTG AAG GCA CTG TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro 660 665 670	2074
GGC CTC CTG GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala 675 680 685	2122
TGG CGC ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro Glu 690 695 700 705	2170
CTG TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro 710 715 720	2218
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG AAC Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln Asn 725 730 735	2266
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC CAT GGG Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His Gly 740 745 750	2314
CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG ACA GAC CTC His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp Leu 755 760 765	2362
CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG GAG ACC AGC CCG Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser Pro 770 775 780 785	2410
CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC TCC CTG AAT GAG GCC Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser Leu Asn Glu Ala 790 795 800	2458
AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC TTC ATG TGC CAC CAC GCC Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met Cys His His Ala 805 810 815	2506
GTG CGC ATC AGG GGC AAG TCC TAC GTC CAG TGC CAG GGG ATC CCG CAG Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln 820 825 830	2554
GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC CTG TGC TAC GGC GAC ATG Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met 835 840 845	2602
GAG AAC AAG CTG TTT GCG GGG ATT CGG CGG GAC GGG CTG CTC CTG CGT Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu Arg 850 855 860	2650

850		855		860		865	
TTG GTG GAT GAT TTC	TTG TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA	2698					
Leu Val Asp Asp Phe	Leu Leu Val Thr Pro His Leu Thr His Ala Lys						
	870 875 880						
ACC TTC CTC AGG ACC CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG	2746						
Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys Val							
	885 890 895						
GTG AAC TTG CGG AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC	2794						
Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu Ala							
	900 905 910						
CTG GGT GGC ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC	2842						
Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe Pro							
	915 920 925						
TGG TGC GGC CTG CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC	2890						
Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser Asp							
	930 935 940 945						
TAC TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC	2938						
Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe Asn							
	950 955 960						
CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC	2986						
Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly Val							
	965 970 975						
TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC	3034						
Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn Ser							
	980 985 990						
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG GCG	3082						
Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Gln Ala							
	995 1000 1005						
TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG CAA GTT	3130						
Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln Gln Val							
	1010 1015 1020 1025						
TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC TCT GAC ACG GCC TCC	3178						
Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp Thr Ala Ser							
	1030 1035 1040						
CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC GCA GGG ATG TCG CTG GGG	3226						
Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser Leu Gly							
	1045 1050 1055						
GCC AAG GGC GCC GCC GGC CCT CTG CCC TCC GAG GCC GTG CAG TGG CTG	3274						
Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln Trp Leu							
	1060 1065 1070						
TGC CAC CAA GCA TTC CTG CTC AAG CTG ACT CGA CAC CGT GTC ACC TAC	3322						
Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg Val Thr Tyr							
	1075 1080 1085						
GTG CCA CTC CTG GGG TCA CTC AGG ACA GCC CAG ACG CAG CTG AGT CGG	3370						
Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr Gln Leu Ser Arg							
	1090 1095 1100 1105						

AAG CTC CCG GGG ACG ACG CTG ACT GCC CTG GAG GCC GCA GCC AAC CCG	3418
Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala Ala Ala Asn Pro	
1110 1115 1120	
GCA CTG CCC TCA GAC TTC AAG ACC ATC CTG GAC TGATGGCCAC CCGCCACAG	3471
Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp	
1125 1130	
CCAGGCCGAG AGCAGACACC AGCAGCCCTG TCACGCCGGG CTCTACGTCC CAGGGAGGGA	3531
GGGGCGGCCC ACACCCAGGC CCGCACCCTG GGGAGTCTGA GGCCTGAGTG AGTGTTTGGC	3591
CGAGGCCTGC ATGTCCGGCT GAAGGCTGAG TGTCCGGCTG AGGCCTGAGC GAGTGTCCAG	3651
CCAAGGGCTG AGTGTCCAGC ACACCTGCCG TCTTCACTTC CCCACAGGCT GGCGCTCGGC	3711
TCCACCCCAG GGCCAGCTTT TCCTCACCAG GAGCCCGGCT TCCACTCCCC ACATAGGAAT	3771
AGTCCATCCC CAGATTCGCC ATTGTTTACC CCTCGCCCTG CCCTCCTTTG CTTTCCACCC	3831
CCACCATCCA GGTGGAGACC CTGAGAAGGA CCCTGGGAGC TCTGGGAATT TGGAGTGACC	3891
AAAGGTGTGC CCTGTACACA GGCGAGGACC CTGCACCTGG ATGGGGGTCC CTGTGGGTCA	3951
AATTGGGGGG AGGTGCTGTG GGAGTAAAAT ACTGAATATA TGAGTTTTTC AGTTTTGAAA	4011
AAAA	4015

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1132 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser
1				5					10					15	
His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr	Phe	Val	Arg	Arg	Leu	Gly
			20					25					30		
Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly	Asp	Pro	Ala	Ala	Phe	Arg
		35					40					45			
Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro
	50					55					60				
Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Leu
65					70					75				80	
Val	Ala	Arg	Val	Leu	Gln	Arg	Leu	Cys	Glu	Arg	Gly	Ala	Lys	Asn	Val
			85					90						95	
Leu	Ala	Phe	Gly	Phe	Ala	Leu	Leu	Asp	Gly	Ala	Arg	Gly	Gly	Pro	Pro
		100						105					110		
Glu	Ala	Phe	Thr	Thr	Ser	Val	Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Thr

115					120					125					
Asp	Ala	Leu	Arg	Gly	Ser	Gly	Ala	Trp	Gly	Leu	Leu	Leu	Arg	Arg	Val
130						135					140				
Gly	Asp	Asp	Val	Leu	Val	His	Leu	Leu	Ala	Arg	Cys	Ala	Leu	Phe	Val
145					150					155					160
Leu	Val	Ala	Pro	Ser	Cys	Ala	Tyr	Gln	Val	Cys	Gly	Pro	Pro	Leu	Tyr
				165					170					175	
Gln	Leu	Gly	Ala	Ala	Thr	Gln	Ala	Arg	Pro	Pro	Pro	His	Ala	Ser	Gly
			180					185					190		
Pro	Arg	Arg	Arg	Leu	Gly	Cys	Glu	Arg	Ala	Trp	Asn	His	Ser	Val	Arg
			195				200					205			
Glu	Ala	Gly	Val	Pro	Leu	Gly	Leu	Pro	Ala	Pro	Gly	Ala	Arg	Arg	Arg
	210					215					220				
Gly	Gly	Ser	Ala	Ser	Arg	Ser	Leu	Pro	Leu	Pro	Lys	Arg	Pro	Arg	Arg
225					230					235					240
Gly	Ala	Ala	Pro	Glu	Pro	Glu	Arg	Thr	Pro	Val	Gly	Gln	Gly	Ser	Trp
				245					250					255	
Ala	His	Pro	Gly	Arg	Thr	Arg	Gly	Pro	Ser	Asp	Arg	Gly	Phe	Cys	Val
			260					265					270		
Val	Ser	Pro	Ala	Arg	Pro	Ala	Glu	Glu	Ala	Thr	Ser	Leu	Glu	Gly	Ala
			275				280					285			
Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg	Gln	His	His
			290			295					300				
Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp	Asp	Thr	Pro
305					310					315					320
Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr	Ser	Ser	Gly
				325					330					335	
Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser	Leu	Arg	Pro
			340					345					350		
Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe	Leu	Gly	Ser
			355				360					365			
Arg	Pro	Trp	Met	Pro	Gly	Thr	Pro	Arg	Arg	Leu	Pro	Arg	Leu	Pro	Gln
	370					375					380				
Arg	Tyr	Trp	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu	Leu	Gly	Asn	His
385					390					395					400
Ala	Gln	Cys	Pro	Tyr	Gly	Val	Leu	Leu	Lys	Thr	His	Cys	Pro	Leu	Arg
				405					410					415	
Ala	Ala	Val	Thr	Pro	Ala	Ala	Gly	Val	Cys	Ala	Arg	Glu	Lys	Pro	Gln
			420				425					430			
Gly	Ser	Val	Ala	Ala	Pro	Glu	Glu	Glu	Asp	Thr	Asp	Pro	Arg	Arg	Leu
			435			440						445			

Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe  
 450 455 460  
 Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser  
 465 470 475 480  
 Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser  
 485 490 495  
 Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met  
 500 505  
 Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys  
 515 520 525  
 Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe  
 530 535 540  
 Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe  
 545 550 555 560  
 Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr  
 565 570 575  
 Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His  
 580 585 590  
 Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln  
 595 600 605  
 His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile  
 610 615 620  
 Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val  
 625 630 635 640  
 Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser  
 645 650 655  
 Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg  
 660 665 670  
 Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg  
 675 680 685  
 Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro  
 690 695 700  
 Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile  
 705 710 715 720  
 Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln  
 725 730 735  
 Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His  
 740 745 750  
 Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp  
 755 760 765  
 Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser  
 770 775 780

Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser Leu Asn Glu  
 785 790 795 800  
 Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met Cys His His  
 805 810 815  
 Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro  
 820 825 830  
 Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp  
 835 840 845  
 Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu  
 850 855 860  
 Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala  
 865 870 875 880  
 Lys Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys  
 885 890 895  
 Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu  
 900 905 910  
 Ala Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe  
 915 920 925  
 Pro Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser  
 930 935 940  
 Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe  
 945 950 955 960  
 Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly  
 965 970 975  
 Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn  
 980 985 990  
 Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln  
 995 1000 1005  
 Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln Gln  
 1010 1015 1020  
 Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp Thr Ala  
 1025 1030 1035 1040  
 Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser Leu  
 1045 1050 1055  
 Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln Trp  
 1060 1065 1070  
 Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg Val Thr  
 1075 1080 1085  
 Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr Gln Leu Ser  
 1090 1095 1100  
 Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala Ala Ala Asn  
 1105 1110 1115 1120

Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp  
 1125 1130

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2176 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..2176
- (D) OTHER INFORMATION: /note= "clone 712562"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCCAAGTTC CTGCACTGGC TGATGAGTGT GTACGTCGTC GAGCTGCTCA GGTCTTTCTT	60
TTATGTCACG GAGACCACGT TTCAAAGAA CAGGCTCTTT TTCTACCGGA AGAGTGTCTG	120
GAGCAAGTTG CAAAGCATTG GAATCAGACA GCACTTGAAG AGGGTGCAGC TCGGGGAGCT	180
GTCGGAAGCA GAGGTCAGGC AGCATCGGGA AGCCAGGCCG GCCCTGCTGA CGTCCAGACT	240
CCGCTTCATC CCCAAGCCTG ACGGGCTGCG GCCGATTGTG AACATGGACT ACGTCGTGGG	300
AGCCAGAACG TTCCGCAGAG AAAAGAGGGC CGAGCGTCTC ACCTCGAGGG TGAAGGCACT	360
GTTTCAGCGTG CTCAACTACG AGCGGGCGCG GCGCCCCGCG CTCCTGGGCG CCTCTGTGCT	420
GGGCCTGGAC GATATCCACA GGGCCTGGCG CACCTTCGTG CTGCGTGTGC GGGCCCAGGA	480
CCCGCCGCCT GAGCTGTACT TTGTCAAGGT GGATGTGACG GGCGCGTACG ACACCATCCC	540
CCAGGACAGG CTCACGGAGG TCATCGCCAG CATCATCAAA CCCCAGAACA CGTACTGCGT	600
GCGTCGGTAT GCCGTGGTCC AGAAGGCCGC CCATGGGCAC GTCCGCAAGG CCTTCAAGAG	660
CCACGTCCTA CGTCCAGTGC CAGGGGATCC CGCAGGGCTC CATCCTCTCC ACGCTGCTCT	720
GCAGCCTGTG CTACGGCGAC ATGGAGAACA AGCTGTTTGC GGGGATTTCG CGGGACGGGC	780
TGCTCCTGCG TTTGGTGGAT GATTTCCTGT TGGTGACACC TCACCTCACC CACGCGAAAA	840
CCTTCCTCAG GACCCTGGTC CGAGGTGTCC CTGAGTATGG CTGCGTGGTG AACTTGCGGA	900
AGACAGTGGT GAACTTCCCT GTAGAAGACG AGGCCCTGGG TGGCACGGCT TTTGTTTACA	960
TGCCGGCCCA CGGCCTATTC CCCTGGTGCG GCCTGCTGCT GGATACCCGG ACCCTGGAGG	1020
TGCAGAGCGA CTAATCCAGC TATGCCCCGA CCTCCATCAG AGCCAGTCTC ACCTTCAACC	1080
GCGGCTTCAA GGCTGGGAGG AACATGCGTC GCAAACTCTT TGGGGTCTTG CGGCTGAAGT	1140
GTACACGCCT GTTTCTGGAT TTGCAGGTGA ACAGCCTCCA GACGGTGTGC ACCAACATCT	1200

ACAAGATCCT CCTGCTGCAG GCGTACAGGT TTCACGCATG TGTGCTGCAG CTCCCATTTC	1260
ATCAGCAAGT TTGGAAGAAC CCCACATTTT TCCTGCGCGT CATCTCTGAC ACGGCCTCCC	1320
TCTGCTACTC CATCCTGAAA GCCAAGAACG CAGGGATGTC GCTGGGGGCC AAGGGCGCCG	1380
CCGGCCNTCT GCCCTCCGAG GCCGTGCAGT GGCTGTGCCA CCAAGCATTG CTGCTCAAGC	1440
TGACTCGACA CCGTGTCAAC TACGTGCCAC TCCTGGGGTC ACTCAGGACA GCCCAGACGC	1500
AGCTGAGTCG GAAGCTCCCG GGGACGACGC TGAAGTCCCT GGAGGCCGCA GCCAACCCGG	1560
CACTGCCCTC AGACTTCAAG ACCATCCTGG ACTGATGGCC ACCCGCCCAC AGCCAGGCCG	1620
AGAGCAGACA CCAGCAGCCC TGTCACGCCG GGCTCTACGT CCCAGGGAGG GAGGGGCGGC	1680
CCACACCCAG GCCTGCACCG CTGGGAGTCT GAGGCCTGAG TGAGTGTTTG GCCGAGGCCT	1740
GCATGTCCCG CTGAAGGCTG AGTGTCCGGC TGAGGCCTGA GCGAGTGTCC AGCCAAGGGC	1800
TGAGTGTTCA GCACACCTGC CGTCTTCACT TCCCCACAGG CTGGCGCTCG GCTCCACCCC	1860
AGGGCCAGCT TTTCCTCACC AGGAGCCCGG CTTCCACTCC CCACATAGGA ATAGTCCATC	1920
CCCAGATTCG CCATTGTTCA CCCCTCGCCC TGCCCTCCTT TGCCTTCCAC CCCCACCATC	1980
CAGGTGGAGA CCCTGAGAAG GACCCTGGGA GCTCTGGGAA TTTGGAGTGA CCAAAGGTGT	2040
GCCCTGTACA CAGGCGAGGA CCCTGCACCT GGATGGGGGT CCCTGTGGGT CAAATTGGGG	2100
GGAGGTGCTG TGGGAGTAAA ATACTGAATA TATGAGTTTT TCAGTTTTGN AAAAAAAAAA	2160
AAAAAAAAAA AAAAAA	2176

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3855 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..3855
- (D) OTHER INFORMATION: /note= "nucleic acid sequence with an open reading frame encoding a delta-182 variant polypeptide"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 56..2479
- (D) OTHER INFORMATION: /product= "delta-182 variant polypeptide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:



GCAGCGCTGC	GTCCTGCTGC	GCACGTGGGA	AGCCCTGGCC	CCGGCCACCC	CCGCG	ATG	58
						Met	
						1	
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC CAC	106						
Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser His							
5 10 15							
TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG GGG CCC	154						
Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly Pro							
20 25 30							
CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT TTC CGC GCG	202						
Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg Ala							
35 40 45							
CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC GCA CGG CCG CCC	250						
Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro Pro							
50 55 60 65							
CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC CTG AAG GAG CTG GTG	298						
Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu Val							
70 75 80							
GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC GGC GCG AAG AAC GTG CTG	346						
Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val Leu							
85 90 95							
GCC TTC GGC TTC GCG CTG CTG GAC GGG GCC CGC GGG GGC CCC CCC GAG	394						
Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Gly Pro Pro Glu							
100 105 110							
GCC TTC ACC ACC AGC GTG CGC AGC TAC CTG CCC AAC ACG GTG ACC GAC	442						
Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr Asp							
115 120 125							
GCA CTG CGG GGG AGC GGG GCG TGG GGG CTG CTG CTG CGC CGC GTG GGC	490						
Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val Gly							
130 135 140 145							
GAC GAC GTG CTG GTT CAC CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG	538						
Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val Leu							
150 155 160							
GTG GCT CCC AGC TGC GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG	586						
Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr Gln							
165 170 175							
CTC GGC GCT GCC ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC	634						
Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly Pro							
180 185 190							
CGA AGG CGT CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG	682						
Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg Glu							
195 200 205							
GCC GGG GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG	730						
Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg Gly							
210 215 220 225							
GGC AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC	778						
Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg Gly							

230										235					240					
GCT Ala	GCC Ala	CCT Pro	GAG Glu 245	CCG Pro	GAG Glu	CGG Arg	ACG Thr	CCC Pro 250	GTT Val	GGG Gly	CAG Gln	GGG Gly	TCC Ser 255	TGG Trp	GCC Ala	826				
CAC His	CCG Pro	GGC Gly 260	AGG Arg	ACG Thr	CGT Arg	GGA Gly	CCG Pro 265	AGT Ser	GAC Asp	CGT Arg	GGT Gly	TTC Phe 270	TGT Cys	GTG Val	GTG Val	874				
TCA Ser	CCT Pro 275	GCC Ala	AGA Arg	CCC Pro	GCC Ala	GAA Glu 280	GAA Glu	GCC Ala	ACC Thr	TCT Ser	TTG Leu 285	GAG Glu	GGT Gly	GCG Ala	CTC Leu	922				
TCT Ser 290	GGC Gly	ACG Thr	CGC Arg	CAC His	TCC Ser 295	CAC His	CCA Pro	TCC Ser	GTG Val	GGC Gly 300	CGC Arg	CAG Gln	CAC His	CAC His	GCG Ala 305	970				
GGC Gly	CCC Pro	CCA Pro	TCC Ser	ACA Thr 310	TCG Ser	CGG Arg	CCA Pro	CCA Pro	CGT Arg 315	CCC Pro	TGG Trp	GAC Asp	ACG Thr	CCT Pro 320	TGT Cys	1018				
CCC Pro	CCG Pro	GTG Val	TAC Tyr 325	GCC Ala	GAG Glu	ACC Thr	AAG Lys	CAC His 330	TTC Phe	CTC Leu	TAC Tyr	TCC Ser	TCA Ser 335	GGC Gly	GAC Asp	1066				
AAG Lys	GAG Glu	CAG Gln 340	CTG Leu	CGG Arg	CCC Pro	TCC Ser	TTC Phe 345	CTA Leu	CTC Leu	AGC Ser	TCT Ser	CTG Leu 350	AGG Arg	CCC Pro	AGC Ser	1114				
CTG Leu 355	ACT Thr	GGC Gly	GCT Ala	CGG Arg	AGG Arg	CTC Leu 360	GTG Val	GAG Glu	ACC Thr	ATC Ile	TTT Phe 365	CTG Leu	GGT Gly	TCC Ser	AGG Arg	1162				
CCC Pro 370	TGG Trp	ATG Met	CCA Pro	GGG Gly 375	ACT Thr	CCC Pro	CGC Arg	AGG Arg	TTG Leu 380	CCC Pro	CGC Arg	CTG Leu	CCC Pro	CAG Gln	CGC Arg 385	1210				
TAC Tyr	TGG Trp	CAA Gln	ATG Met	CGG Arg 390	CCC Pro	CTG Leu	TTT Phe	CTG Leu	GAG Glu 395	CTG Leu	CTT Leu	GGG Gly	AAC Asn	CAC His 400	GCG Ala	1258				
CAG Gln	TGC Cys	CCC Pro	TAC Tyr 405	GGG Gly	GTG Val	CTC Leu	CTC Leu	AAG Lys 410	ACG Thr	CAC His	TGC Cys	CCG Pro	CTG Leu 415	CGA Arg	GCT Ala	1306				
GCG Ala	GTC Val	ACC Thr 420	CCA Pro	GCA Ala	GCC Ala	GGT Gly 425	GTC Val	TGT Cys	GCC Ala	CGG Arg	GAG Glu	AAG Lys 430	CCC Pro	CAG Gln	GGC Gly	1354				
TCT Ser	GTG Val 435	GCG Ala	GCC Ala	CCC Pro	GAG Glu	GAG Glu 440	GAG Glu	GAC Asp	ACA Thr	GAC Asp	CCC Pro 445	CGT Arg	CGC Arg	CTG Leu	GTG Val	1402				
CAG Gln 450	CTG Leu	CTC Leu	CGC Arg	CAG Gln	CAC His 455	AGC Ser	AGC Ser	CCC Pro	TGG Trp	CAG Gln 460	GTG Val	TAC Tyr	GGC Gly	TTC Phe	GTG Val 465	1450				
CGG Arg	GCC Ala	TGC Cys	CTG Leu	CGC Arg 470	CGG Arg	CTG Leu	GTG Val	CCC Pro 475	CCA Pro	GGC Gly	CTC Leu	TGG Trp	GGC Gly	TCC Ser 480	AGG Arg	1498				

CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC CTG His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser Leu 485 490 495	1546
GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG ATG AGC Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met Ser 500 505 510	1594
GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT GGC TGT GTT Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys Val 515 520 525	1642
CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG GCC AAG TTC CTG Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe Leu 530 535 540 545	1690
CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG CTC AGG TCT TTC TTT His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe 550 555 560	1738
TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG CTC TTT TTC TAC CGG Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg 565 570 575	1786
AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT GGA ATC AGA CAG CAC TTG Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu 580 585 590	1834
AAG AGG GTG CAG CTG CGG GAG CTG TCG GAA GCA GAG GTC AGG CAG CAT Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln His 595 600 605	1882
CGG GAA GCC AGG CCC GCC CTG CTG ACG TCC AGA CTC CGC TTC ATC CCC Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro 610 615 620 625	1930
AAG CCT GAC GGG CTG CGG CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val Gly 630 635 640	1978
GCC AGA ACG TTC CGC AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg 645 650 655	2026
GTG AAG GCA CTG TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro 660 665 670	2074
GGC CTC CTG GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala 675 680 685	2122
TGG CGC ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro Glu 690 695 700 705	2170
CTG TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro 710 715 720	2218
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG AAC Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln Asn	2266

725										730					735					
ACG	TAC	TGC	GTG	CGT	CGG	TAT	GCC	GTG	GTC	CAG	AAG	GCC	GCC	CAT	GGG	2314				
Thr	Tyr	Cys	Val	Arg	Arg	Tyr	Ala	Val	Val	Gln	Lys	Ala	Ala	His	Gly					
		740					745					750								
CAC	GTC	CGC	AAG	GCC	TTC	AAG	AGC	CAC	GTC	CTA	CGT	CCA	GTG	CCA	GGG	2362				
His	Val	Arg	Lys	Ala	Phe	Lys	Ser	His	Val	Leu	Arg	Pro	Val	Pro	Gly					
	755					760					765									
GAT	CCC	GCA	GGG	CTC	CAT	CCT	CTC	CAC	GCT	GCT	CTG	CAG	CCT	GTG	CTA	2410				
Asp	Pro	Ala	Gly	Leu	His	Pro	Leu	His	Ala	Ala	Leu	Gln	Pro	Val	Leu					
	770				775				780						785					
CGG	CGA	CAT	GGA	GAA	CAA	GCT	GTT	TGC	GGG	GAT	TCG	GCG	GGA	CGG	GCT	2458				
Arg	Arg	His	Gly	Glu	Gln	Ala	Val	Cys	Gly	Asp	Ser	Ala	Gly	Arg	Ala					
			790						795					800						
GCT	CCT	GCG	TTT	GGT	GGA	TGATTTCTTG	TTGGTGACAC	CTCACCTCAC								2506				
Ala	Pro	Ala	Phe	Gly	Gly															
			805																	
CCACGCGAAA	ACCTTCCTCA	GGACCCTGGT	CCGAGGTGTC	CCTGAGTATG	GCTGCGTGGT											2566				
GAACCTTGCGG	AAGACAGTGG	TGAACTTCCC	TGTAGAAGAC	GAGGCCCTGG	GTGGCACGGC											2626				
TTTTGTTCAG	ATGCCGGCCC	ACGGCCTATT	CCCCTGGTGC	GGCCTGCTGC	TGGATACCCG											2686				
GACCCTGGAG	GTGCAGAGCG	ACTACTCCAG	CTATGCCCCG	ACCTCCATCA	GAGCCAGTCT											2746				
CACCTTCAAC	CGCGGCTTCA	AGGCTGGGAG	GAACATGCGT	CGCAAACCTCT	TTGGGGTCTT											2806				
GCGGCTGAAG	TGTCACAGCC	TGTTTCTGGA	TTTGCAGGTG	AACAGCCTCC	AGACGGTGTG											2866				
CACCAACATC	TACAAGATCC	TCCTGCTGCA	GGCGTACAGG	TTTCACGCAT	GTGTGCTGCA											2926				
GCTCCCATTT	CATCAGCAAG	TTTGGAAGAA	CCCCACATTT	TTCCTGCGCG	TCATCTCTGA											2986				
CACGGCCTCC	CTCTGCTACT	CCATCCTGAA	AGCCAAGAAC	GCAGGGATGT	CGCTGGGGGC											3046				
CAAGGGCGCC	GCCGGCCCTC	TGCCCTCCGA	GGCCGTGCAG	TGGCTGTGCC	ACCAAGCATT											3106				
CCTGCTCAAG	CTGACTCGAC	ACCGTGTAC	CTACGTGCCA	CTCCTGGGGT	CACTCAGGAC											3166				
AGCCCAGACG	CAGCTGAGTC	GGAAGCTCCC	GGGGACGACG	CTGACTGCCC	TGGAGGCCGC											3226				
AGCCAACCCG	GCACTGCCCT	CAGACTTCAA	GACCATCCTG	GACTGATGGC	CACCCGCCCA											3286				
CAGCCAGGCC	GAGAGCAGAC	ACCAGCAGCC	CTGTACGCC	GGGCTCTACG	TCCCAGGGAG											3346				
GGAGGGGCGG	CCCACACCCA	GGCCCGCACC	GCTGGGAGTC	TGAGGCCTGA	GTGAGTGTTT											3406				
GGCCGAGGCC	TGCATGTCCG	GCTGAAGGCT	GAGTGTCCGG	CTGAGGCCTG	AGCGAGTGTC											3466				
CAGCCAAGGG	CTGAGTGTCC	AGCACACCTG	CCGTCTTCAC	TTCCCCACAG	GCTGGCGCTC											3526				
GGCTCCACCC	CAGGGCCAGC	TTTTCCCTCAC	CAGGAGCCCCG	GCTTCCACTC	CCCACATAGG											3586				
AATAGTCCAT	CCCCAGATTC	GCCATTGTTC	ACCCCTCGCC	CTGCCCTCCT	TTGCCCTTCCA											3646				
CCCCCACCAT	CCAGGTGGAG	ACCCTGAGAA	GGACCCTGGG	AGCTCTGGGA	ATTTGGAGTG											3706				

ACCAAAGGTG TGCCCTGTAC ACAGGCGAGG ACCCTGCACC TGGATGGGGG TCCCTGTGGG	3766
TCAAATTGGG GGGAGGTGCT GTGGGAGTAA AATACTGAAT ATATGAGTTT TTCAGTTTTG	3826
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA	3855

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 807 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser	
1				5					10					15		
His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr	Phe	Val	Arg	Arg	Leu	Gly	
			20					25						30		
Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly	Asp	Pro	Ala	Ala	Phe	Arg	
		35					40					45				
Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro	
	50					55					60					
Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Leu	
65					70					75				80		
Val	Ala	Arg	Val	Leu	Gln	Arg	Leu	Cys	Glu	Arg	Gly	Ala	Lys	Asn	Val	
			85					90						95		
Leu	Ala	Phe	Gly	Phe	Ala	Leu	Leu	Asp	Gly	Ala	Arg	Gly	Gly	Pro	Pro	
		100						105					110			
Glu	Ala	Phe	Thr	Thr	Ser	Val	Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Thr	
		115					120					125				
Asp	Ala	Leu	Arg	Gly	Ser	Gly	Ala	Trp	Gly	Leu	Leu	Leu	Arg	Arg	Val	
	130					135					140					
Gly	Asp	Asp	Val	Leu	Val	His	Leu	Leu	Ala	Arg	Cys	Ala	Leu	Phe	Val	
145				150					155					160		
Leu	Val	Ala	Pro	Ser	Cys	Ala	Tyr	Gln	Val	Cys	Gly	Pro	Pro	Leu	Tyr	
			165					170						175		
Gln	Leu	Gly	Ala	Ala	Thr	Gln	Ala	Arg	Pro	Pro	Pro	His	Ala	Ser	Gly	
		180					185						190			
Pro	Arg	Arg	Arg	Leu	Gly	Cys	Glu	Arg	Ala	Trp	Asn	His	Ser	Val	Arg	
		195					200					205				
Glu	Ala	Gly	Val	Pro	Leu	Gly	Leu	Pro	Ala	Pro	Gly	Ala	Arg	Arg	Arg	
	210					215					220					
Gly	Gly	Ser	Ala	Ser	Arg	Ser	Leu	Pro	Leu	Pro	Lys	Arg	Pro	Arg	Arg	
225					230					235				240		

Gly	Ala	Ala	Pro	Glu	Pro	Glu	Arg	Thr	Pro	Val	Gly	Gln	Gly	Ser	Trp
				245					250					255	
Ala	His	Pro	Gly	Arg	Thr	Arg	Gly	Pro	Ser	Asp	Arg	Gly	Phe	Cys	Val
			260					265					270		
Val	Ser	Pro	Ala	Arg	Pro	Ala	Glu	Glu	Ala	Thr	Ser	Leu	Glu	Gly	Ala
		275					280					285			
Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg	Gln	His	His
	290					295					300				
Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp	Asp	Thr	Pro
305					310					315					320
Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr	Ser	Ser	Gly
				325					330					335	
Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser	Leu	Arg	Pro
			340					345					350		
Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe	Leu	Gly	Ser
		355					360					365			
Arg	Pro	Trp	Met	Pro	Gly	Thr	Pro	Arg	Arg	Leu	Pro	Arg	Leu	Pro	Gln
	370					375					380				
Arg	Tyr	Trp	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu	Leu	Gly	Asn	His
385					390					395					400
Ala	Gln	Cys	Pro	Tyr	Gly	Val	Leu	Leu	Lys	Thr	His	Cys	Pro	Leu	Arg
				405					410					415	
Ala	Ala	Val	Thr	Pro	Ala	Ala	Gly	Val	Cys	Ala	Arg	Glu	Lys	Pro	Gln
			420					425					430		
Gly	Ser	Val	Ala	Ala	Pro	Glu	Glu	Glu	Asp	Thr	Asp	Pro	Arg	Arg	Leu
		435					440					445			
Val	Gln	Leu	Leu	Arg	Gln	His	Ser	Ser	Pro	Trp	Gln	Val	Tyr	Gly	Phe
	450					455					460				
Val	Arg	Ala	Cys	Leu	Arg	Arg	Leu	Val	Pro	Pro	Gly	Leu	Trp	Gly	Ser
465					470				475						480
Arg	His	Asn	Glu	Arg	Arg	Phe	Leu	Arg	Asn	Thr	Lys	Lys	Phe	Ile	Ser
				485					490					495	
Leu	Gly	Lys	His	Ala	Lys	Leu	Ser	Leu	Gln	Glu	Leu	Thr	Trp	Lys	Met
			500					505					510		
Ser	Val	Arg	Asp	Cys	Ala	Trp	Leu	Arg	Arg	Ser	Pro	Gly	Val	Gly	Cys
		515					520					525			
Val	Pro	Ala	Ala	Glu	His	Arg	Leu	Arg	Glu	Glu	Ile	Leu	Ala	Lys	Phe
	530					535					540				
Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu	Arg	Ser	Phe
545					550					555					560
Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu	Phe	Phe	Tyr
				565					570					575	

Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His  
 580 585 590  
 Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln  
 595 600 605  
 His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile  
 610 615 620  
 Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val  
 625 630 635 640  
 Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser  
 645 650 655  
 Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg  
 660 665 670  
 Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg  
 675 680 685  
 Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro  
 690 695 700  
 Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile  
 705 710 715 720  
 Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln  
 725 730 735  
 Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His  
 740 745 750  
 Gly His Val Arg Lys Ala Phe Lys Ser His Val Leu Arg Pro Val Pro  
 755 760 765  
 Gly Asp Pro Ala Gly Leu His Pro Leu His Ala Ala Leu Gln Pro Val  
 770 775 780  
 Leu Arg Arg His Gly Glu Gln Ala Val Cys Gly Asp Ser Ala Gly Arg  
 785 790 795 800  
 Ala Ala Pro Ala Phe Gly Gly  
 805

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATCGATTGGG CCCGAGATCT CGCGCGCGAG GCCTGCCATG GGACCCACTG CAGGGGCAGC

60

TGGGANGCTG	CAGGCTTCAG	GTCCCAGTGG	GGTTGCCATC	TGCCAGTAGA	AACCTGATGT	120
AGAATCAGGG	CGCGAGTGTG	GACACTGTCC	TGAATCTCAA	TGTCTCAGTG	TGTGCTGAAA	180
CATGTAGAAA	TTAAAGTCCA	TCCCTCCTAC	TCTACTGGGA	TTGAGCCCCT	TCCCTATCCC	240
CCCCCAGGGG	CAGAGGAGTT	CCTCTCACTC	CTGTGGAGGA	AGGAATGATA	CTTTGTTATT	300
TTTCACTGCT	GGTACTGAAT	CCACTGTTTC	ATTTGTTGGT	TTGTTTGT	TGTTTTGAGA	360
AGCGGTTTCA	CTCTTGTTGC	TCAGGCTGGA	NGGAGTGCAA	TGGCGCGATC	TTGGCTTACT	420
GCAGCCTCTG	CCTCCCAGGT	TCAAGTGATT	CTCCTGCTTC	CGCCTCCCAT	TTGGCTGGGA	480
TTACAGGCAC	CCGCCACCAT	GCCCAGCTAA	TTTTTTGTAT	TTTTAGTANA	NACNGGGGTG	540
GGGGTGGGGT	TCACATGTTG	GCCAAGCTGG	TCTCGAACTT	CTGAACTCAG	ATGATCCANC	600
TGCCTCTGCC	TCCTAAAATT	GCTGGGATTA	CAGGTGTNAN	CCACCATGCC	CAACTCAAAA	660
TTTACTCTGT	TTANAAACAT	CTGGGTCTAA	GGTAGGAANC	TCACCCCACT	CAATTTTTGT	720
GGTGTTTTTA	AGCCAATNAN	AAAATTTTTT	NATGTTGTTT	NNNNNNNNNN	NNNNNNNNNN	780
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	840
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	900
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	960
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	1020
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	1080
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	1140
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	1200
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	1260
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	1320
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	1380
NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNCCGG	TGNNNGAGGG	1440
NGCCANGRAG	GGGGCCAGGT	TCCAANTTCC	CAACCKTTTT	WGGARGGACN	GCCCCCAGGG	1500
GGGGATRAAC	AGANTNGGGG	GKGGTWGGGT	TNAKGGTGGG	AACNCCTTNG	CSGCCTGGAG	1560
AACGTGCAAA	GAGGAAATGA	AGGGCCTGKG	TCAAGGAGCC	CAAGTNGGCG	GGGRAGTTTG	1620
CAGGGAGGCA	CTCCGGGGAG	GTCCSGCGTG	CCCGTCCAAG	GGAGCAATGC	GTCCTTCGGG	1680
TTCGTCCCCA	WGCGCGTCT	ACGCGCCTYC	CGTCCTCCCC	TTCACGTTCC	GGCATTCTGT	1740
GTGCCCCGAG	CCCGACGCCC	CGCGTCCGGA	CCTGGAGGCA	GCCCTGGGTC	TCCGGATCAG	1800
GCCAGCGGCC	AAAGGGTCGC	CGCACGCACC	TGTTCCCAGG	GCCTCCACAT	CATGGCCCCCT	1860
CCCTCGGGTT	ACCCACAGC	CTAGGCCGGA	TTCGACCTCT	CTCCGCTGGG	GCCCTCGCCT	1920



GGCGTCCCTG	CACCCTGGGA	GCGCGAGCGG	CGCGCGGGCG	GGGAAGCGCG	GCCCATACCC	1980
CCGGGTCCGC	CCGAAGCAG	CTGCGCTGTC	GGGGCCAGGC	CGGGCTCCCA	GTGGATTTCG	2040
GGGCACAGAC	GCCCAGGACC	GCGCTTCCCA	CGTGGCGGAA	GGACTGGGGA	CCCGGGCACC	2100
CGTCCTGCCC	CTTCACCTTC	CAGCTCCGCT	TCTTCCGCGC	GGACCCGGCC	CCGTCCCGAA	2160
CCCTTCCCAG	GTCCCGGCCC	AGCCCCCTTC	GGGCCCTCCC	AGCCCCCTCC	CTTCCTTTTC	2220
CGCGGCCCCG	CCCTCTCCTT	CGCGGCGCGA	GTTTCAGGCA	GCGCTGCGTC	CTGCTGCGCA	2280
CGTGGGAAGC	CCTGGCCCCG	GCCACCCCCG	CGATGCCGCG	CGCTCCCCGC	TGCCGAGCCG	2340
TGCGCTCCCT	GCTGCGCAGC	CACTACCGCG	AGGTGCTGCC	GCTGGCCACG	TTCGTGCGGC	2400
GCCTGGGGCC	CCAGGGCTGG	CGGCTGGTGC	AGCGCGGGGA	CCCGGCGGCT	TTCCGCGCGC	2460
TGGTGGCCCA	GTGCTGGTG	TGCGTGCCCT	GGGACGCACG	GCCGCCCCCC	GCCGCCCCCT	2520
CCTTCCGCCA	GGTGGGCCTC	CCCGGGGTGC	GCGTCCGGCT	GGGGTTGAGG	GCGGCCGGGG	2580
GGAACCAGCG	ACATGCGGAG	AGCAGCGCAG	GCGACTCAGG	GCGCTTCCCC	CGCAGGTGTC	2640
CTGCCTGAAG	GAGCTGGTGG	CCCGAGTGCT	GCAGAGGCTG	TGCGAGCGCG	GCGCGAAGAA	2700
CGTGCTGGCC	TTGGGCTTCG	CGCTGCTGGA	CGGGGCCCCG	GGGGGCCCCC	CCGAGGCCTT	2760
CACCACCAGC	GTGCGCAGCT	ACCTGCCCAA	CACGGTGACC	GACGCACTGC	GGGGGAGCGG	2820
GGCGTGGGGG	CTGCTGCTGC	GCCGCGTGGG	CGACGACGTG	CTGGTTCACC	TGCTGGCACG	2880
CTGCGCGCTC	TTTGTGCTGG	TGGCTCCCAG	CTGCGCCTAC	CAGGTGTGCG	GGCCGCCGCT	2940
GTACCAGCTC	GGCGCTGCCA	CTCAGGCCCC	GCCCCCGCCA	CACGCTAGTG	GACCCCGAAG	3000
GCCTCTGGGA	TGCGAACGGG	CCTGGAACCA	TAGCGTCAGG	GAGGCCGGGG	TCCCCCTGGG	3060
CCTGCCAGCC	CCGGGTGCCA	GGAGGCGCGG	GGGCAGTGCC	AGCCGAAGTC	TGCCGTTGCC	3120
CAAGAGGCCC	AGGCGTGGCG	CTGCCCCCTG	GCCGGAGCGG	ACGCCCCGTT	GGCAGGGGTC	3180
CTGGGCCCCA	CCGGGCAGGA	CGCGTGGAAC	GAGTGACCGT	GGTTTCTGTG	TGGTGTACAC	3240
TGCCAGACCC	GCCGAAGAAG	CCACCTCTTT	GGAGGGTGCG	CTCTCTGGCA	CGCGCCACTC	3300
CCACCCATCC	GTGGGCGGCC	AGCACCACGC	GGGCCCCCCA	TCCACATCGC	GGCCACCACG	3360
TCCCTGGGAC	ACGCCTTGTC	CCCCGGTGTA	CGCCGAGACC	AAGCACTTCC	TCTACTCCTC	3420
AGGCGACAAG	GAGCAGCTGC	GGCCCTCCTT	CCTACTCAGC	TCTCTGAGGC	CCAGCCTGAC	3480
TGGCGCTCGG	AGGCTCGTGG	AGACCATCTT	TCTGGGTTCC	AGGCCCTGGA	TGCCAGGGAC	3540
TCCCCGCAGG	TTGCCCCGCC	TGCCCGAGCG	CTACTGGCAA	ATGCGGCCCC	TGTTTCTGGA	3600
GCTGCTTGGG	AACCACGCGC	AGTGCCCCCT	CGGGGTGCTC	CTCAAGACGC	ACTGCCCCGT	3660
GCGAGCTGCG	GTCACCCAG	CAGCCGGTGT	CTGTGCCCCG	GAGAAGCCCC	AGGGCTCTGT	3720
GGCGGCCCCC	GAGGAGGAGG	ACACAGACCC	CCGTCGCCTG	GTGCAGCTGC	TCCGCCAGCA	3780

CAGCAGCCCC TGGCAGGTGT ACGGCTTCGT GCGGGCCTGC CTGCGCCGGC TGGTGCCCCC	3840
AGGCCTCTGG GGCTCCAGGC ACAACGAACG CCGCTTCCTC AGGAACACCA AGAAGTTCAT	3900
CTCCCTGGGG AAGCATGCCA AGCTCTCGCT GCAGGAGCTG ACGTGGAAGA TGAGCGTGCG	3960
GGACTGCGCT TGGCTGCGCA GGAGCCCAGG TGAGGAGGTG GTGGCCGTCG AGGGCCCAGG	4020
CCCCAGAGCT GAATGCAGTA GGGGCTCAGA AAAGGGGGCA GGCAGAGCCC TGGTCCTCCT	4080
GTCTCCATCG TCACGTGGGC ACACGTGGCT TTTCGCTCAG GACGTCGAGT GGACACGGTG	4140
ATCGAGGTCG ACTCTAGAGG ATCCCCGGGT ACCGAGCTCG AATTCGTAAT CATGGTCATA	4200

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 240 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
  - (A) NAME/KEY: intron
  - (B) LOCATION: 95..198
  - (D) OTHER INFORMATION: /note= "intron 1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGACCCGGCG GCTTTCCGCG CGCTGGTGGC CCAGTGCCTG GTGTGCGTGC CCTGGGACGC	60
ACGGCCGCCC CCCGCCGCC CCTCCTTCG CCAGGTGGGC CTCCCCGGGG TCGGCGTCCG	120
GCTGGGGTTG AGGGCGGCCG GGGGGAACCA GCGACATGCG GAGAGCAGCG CAGGCGACTC	180
AGGGCGCTTC CCCCGCAGGT GTCCTGCCTG AAGGAGCTGG TGGCCCGAGT GCTGCAGAGG	240

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 389 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
  - (A) NAME/KEY: -
  - (B) LOCATION: 1..389
  - (D) OTHER INFORMATION: /note= "expressed sequence tag (EST)  
AA281296"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCCAAGTTCC TGCACTGGCT GATGAGTGTG TACGTCGTCG AGCTGCTCAG GTCTTTCTTT	60
---	----

TATGTCACGG AGACCACGTT TCAAAAGAAC AGGCTCTTTT TCTACCGGAA GAGTGTCTGG	120
AGCAAGTTGC AAAGCATTGG AATCAGACAG CACTTGAAGA GGGTGCAGCT GCGGGACGTG	180
TCGGAAGCAG AGGTCAGGCA GCATCGGGAA GCCAGGCCCG CCCTGCTGAC GTCCAGACTC	240
CGCTTCATCC CCAAGCCTGA CGGGCTGCGG CCGATTGTGA ACATGGACTA CGTCGTGGGA	300
GCCAGAACGT TCCGCAGAGA AAAGAGGGCC GAGCGTCTCA CCTCGAGGGT GAAGGCACTG	360
TTCAGCGTGC TCAACTACGA GCGGGCGCG	389

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..182
- (D) OTHER INFORMATION: /note= "182 basepair sequence deleted in clone 712562"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TCTACCTTGA CAGACCTCCA GCCGTACATG CGACAGTTCG TGGCTCACCT GCAGGAGACC	60
AGCCCGCTGA GGGATGCCGT CGTCATCGAG CAGAGCTCCT CCCTGAATGA GGCCAGCAGT	120
GGCCTCTTCG ACGTCTTCCT ACGCTTCATG TGCCACCACG CCGTGCGCAT CAGGGGCAAG	180
TC	182

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 259 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION: 1..259
- (D) OTHER INFORMATION: /note= "protein encoded by clone 712562"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr

1	5	10	15
Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val	20	25	30
Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu Lys Arg Val	35	40	45
Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln His Arg Glu Ala	50	55	60
Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp	65	70	80
Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val Gly Ala Arg Thr	85	90	95
Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg Val Lys Ala	100	105	110
Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu	115	120	125
Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala Trp Arg Thr	130	135	140
Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro Glu Leu Tyr Phe	145	150	155
Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro Gln Asp Arg	165	170	175
Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys	180	185	190
Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His Gly His Val Arg	195	200	205
Lys Ala Phe Lys Ser His Val Leu Arg Pro Val Pro Gly Asp Pro Ala	210	215	220
Gly Leu His Pro Leu His Ala Ala Leu Gln Pro Val Leu Arg Arg His	225	230	235
Gly Glu Gln Ala Val Cys Gly Asp Ser Ala Gly Arg Ala Ala Pro Ala	245	250	255
Phe Gly Gly			

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 34 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Modified-site  
(B) LOCATION: 2  
(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Leu or Ile"

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 10  
(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Leu or Ile"

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 11  
(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Leu or Ile"

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 12  
(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Gln or Arg"

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 28  
(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Phe or Tyr"

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 29  
(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Phe or Tyr"

(ix) FEATURE:  
(A) NAME/KEY: Modified-site  
(B) LOCATION: 31  
(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Lys or His"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Trp	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Phe	Phe	Trp
1				5					10						15	
Xaa	Thr	Glu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Arg	Xaa	Xaa
			20					25						30		
Xaa	Trp															

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 2  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = Leu or Ile"

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 10  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = Leu or Ile"

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 11  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = Leu or Ile"

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 12  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = Gln or Arg"

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 29  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = Phe or Tyr"

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 30  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = Phe or Tyr"

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 32  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = Lys or His"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Trp	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Phe	Phe	Trp
1				5					10					15	
Xaa	Thr	Glu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Arg	Xaa
			20					25					30		
Xaa	Xaa	Trp													
		35													

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 129 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..129
- (D) OTHER INFORMATION: /note= "TRT motifs from human"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ala	Lys	Phe	Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu
1				5					10					15	
Arg	Ser	Phe	Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu
			20				25						30		
Phe	Phe	Tyr	Arg	Lys	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile
		35				40					45				
Arg	Gln	His	Leu	Lys	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser	Glu	Ala	Glu
	50					55					60				
Val	Arg	Gln	His	Arg	Glu	Ala	Arg	Pro	Ala	Leu	Leu	Thr	Ser	Arg	Leu
65				70					75					80	
Arg	Phe	Ile	Pro	Lys	Pro	Asp	Gly	Leu	Arg	Pro	Ile	Val	Asn	Met	Asp
			85					90					95		
Tyr	Val	Val	Gly	Ala	Arg	Thr	Phe	Arg	Arg	Glu	Lys	Arg	Ala	Glu	Arg
			100				105						110		
Leu	Thr	Ser	Arg	Val	Lys	Ala	Leu	Phe	Ser	Val	Leu	Asn	Tyr	Glu	Arg
		115					120					125			

Ala

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..233
- (D) OTHER INFORMATION: /note= "TRT motifs from Schizosaccharomyces pombe tez1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ile	Ser	Glu	Ile	Glu	Trp	Leu	Val	Leu	Gly	Lys	Arg	Ser	Asn	Ala	Lys
1				5					10					15	

Met Cys Leu Ser Asp Phe Glu Lys Arg Lys Gln Ile Phe Ala Glu Phe  
 20 25 30  
 Ile Tyr Trp Leu Tyr Asn Ser Phe Ile Ile Pro Ile Leu Gln Ser Phe  
 35 40 45  
 Phe Tyr Ile Thr Glu Ser Ser Asp Leu Arg Asn Arg Thr Val Tyr Phe  
 50 55 60  
 Arg Lys Asp Ile Trp Lys Leu Leu Cys Arg Pro Phe Ile Thr Ser Met  
 65 70 75 80  
 Lys Met Glu Ala Phe Glu Lys Ile Asn Glu Asn Asn Val Arg Met Asp  
 85 90 95  
 Thr Gln Lys Thr Thr Leu Pro Pro Ala Val Ile Arg Leu Leu Pro Lys  
 100 105 110  
 Lys Asn Thr Phe Arg Leu Ile Thr Asn Leu Arg Lys Arg Phe Leu Ile  
 115 120 125  
 Lys Met Gly Ser Asn Lys Lys Met Leu Val Ser Thr Asn Gln Thr Leu  
 130 135 140  
 Arg Pro Val Ala Ser Ile Leu Lys His Leu Ile Asn Glu Glu Ser Ser  
 145 150 155 160  
 Gly Ile Pro Phe Asn Leu Glu Val Tyr Met Lys Leu Leu Thr Phe Lys  
 165 170 175  
 Lys Asp Leu Leu Lys His Arg Met Phe Gly Arg Lys Lys Tyr Phe Val  
 180 185 190  
 Arg Ile Asp Ile Lys Ser Cys Tyr Asp Arg Ile Lys Gln Asp Leu Met  
 195 200 205  
 Phe Arg Ile Val Lys Lys Lys Leu Lys Asp Pro Glu Phe Val Ile Arg  
 210 215 220  
 Lys Tyr Ala Thr Ile His Ala Thr Ser  
 225 230

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..233
- (D) OTHER INFORMATION: /note= "TRT motifs from Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:



Leu	Lys	Asp	Phe	Arg	Trp	Leu	Phe	Ile	Ser	Asp	Ile	Trp	Phe	Thr	Lys	1	5	10	15
His	Asn	Phe	Glu	Asn	Leu	Asn	Gln	Leu	Ala	Ile	Cys	Phe	Ile	Ser	Trp	20	25	30	
Leu	Phe	Arg	Gln	Leu	Ile	Pro	Lys	Ile	Ile	Gln	Thr	Phe	Phe	Tyr	Cys	35	40	45	
Thr	Glu	Ile	Ser	Ser	Thr	Val	Thr	Ile	Val	Tyr	Phe	Arg	His	Asp	Thr	50	55	60	
Trp	Asn	Lys	Leu	Ile	Thr	Pro	Phe	Ile	Val	Glu	Tyr	Phe	Lys	Thr	Tyr	65	70	75	80
Leu	Val	Glu	Asn	Asn	Val	Cys	Arg	Asn	His	Asn	Ser	Tyr	Thr	Leu	Ser	85	90	95	
Asn	Phe	Asn	His	Ser	Lys	Met	Arg	Ile	Ile	Pro	Lys	Lys	Ser	Asn	Asn	100	105	110	
Glu	Phe	Arg	Ile	Ile	Ala	Ile	Pro	Cys	Arg	Gly	Ala	Asp	Glu	Glu	Glu	115	120	125	
Phe	Thr	Ile	Tyr	Lys	Glu	Asn	His	Lys	Asn	Ala	Ile	Gln	Pro	Thr	Gln	130	135	140	
Lys	Ile	Leu	Glu	Tyr	Leu	Arg	Asn	Lys	Arg	Pro	Thr	Ser	Phe	Thr	Lys	145	150	155	160
Ile	Tyr	Ser	Pro	Thr	Gln	Ile	Ala	Asp	Arg	Ile	Lys	Glu	Phe	Lys	Gln	165	170	175	
Arg	Leu	Leu	Lys	Lys	Phe	Asn	Asn	Val	Leu	Pro	Glu	Leu	Tyr	Phe	Met	180	185	190	
Lys	Phe	Asp	Val	Lys	Ser	Cys	Tyr	Asp	Ser	Ile	Pro	Arg	Met	Glu	Cys	195	200	205	
Met	Arg	Ile	Leu	Lys	Asp	Ala	Leu	Lys	Asn	Glu	Asn	Gly	Phe	Phe	Val	210	215	220	
Arg	Ser	Gln	Tyr	Phe	Phe	Asn	Thr	Asn	225	230									

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..233
- (D) OTHER INFORMATION: /note= "TRT motifs from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Thr	Arg	Glu	Ile	Ser	Trp	Met	Gln	Val	Glu	Thr	Ser	Ala	Lys	His	Phe	
1				5					10					15		
Tyr	Tyr	Phe	Asp	His	Glu	Asn	Ile	Tyr	Val	Leu	Trp	Lys	Leu	Leu	Arg	
		20						25					30			
Trp	Ile	Phe	Glu	Asp	Leu	Val	Val	Ser	Leu	Ile	Arg	Cys	Phe	Phe	Tyr	
		35					40					45				
Val	Thr	Glu	Gln	Gln	Lys	Ser	Tyr	Ser	Lys	Thr	Tyr	Tyr	Tyr	Arg	Lys	
	50					55					60					
Asn	Ile	Trp	Asp	Val	Ile	Met	Lys	Met	Ser	Ile	Ala	Asp	Leu	Lys	Lys	
65					70					75					80	
Glu	Thr	Leu	Ala	Glu	Val	Gln	Glu	Lys	Glu	Val	Glu	Glu	Trp	Lys	Lys	
				85					90					95		
Ser	Leu	Gly	Phe	Ala	Pro	Gly	Lys	Leu	Arg	Leu	Ile	Pro	Lys	Lys	Thr	
			100					105						110		
Thr	Phe	Arg	Pro	Ile	Met	Thr	Phe	Asn	Lys	Lys	Ile	Val	Asn	Ser	Asp	
		115					120					125				
Arg	Lys	Thr	Thr	Lys	Leu	Thr	Thr	Asn	Thr	Lys	Leu	Leu	Asn	Ser	His	
	130					135					140					
Leu	Met	Leu	Lys	Thr	Leu	Lys	Asn	Arg	Met	Phe	Lys	Asp	Pro	Phe	Gly	
145					150					155					160	
Phe	Ala	Val	Phe	Asn	Tyr	Asp	Asp	Val	Met	Lys	Lys	Tyr	Glu	Glu	Phe	
			165						170					175		
Val	Cys	Lys	Trp	Lys	Gln	Val	Gly	Gln	Pro	Lys	Leu	Phe	Phe	Ala	Thr	
			180					185					190			
Met	Asp	Ile	Glu	Lys	Cys	Tyr	Asp	Ser	Val	Asn	Arg	Glu	Lys	Leu	Ser	
		195					200					205				
Thr	Phe	Leu	Lys	Thr	Thr	Lys	Leu	Leu	Ser	Ser	Asp	Phe	Trp	Ile	Met	
	210					215					220					
Thr	Ala	Gln	Ile	Leu	Lys	Arg	Lys	Asn								
225						230										

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:
- (A) NAME/KEY: Peptide

- (B) LOCATION: 1..4
- (D) OTHER INFORMATION: /note= "consensus telomerase RT sequence from motif T"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = polar amino acid, Gly, Ser, Thr, Tyr, Cys, Asn or Gln"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Xaa Phe Phe Tyr  
1

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..7
- (D) OTHER INFORMATION: /note= "consensus telomerase RT sequence from motif 1"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid, Ala, Leu, Ile, Val, Pro, Phe, Trp or Met"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 3
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid, Ala, Leu, Ile, Val, Pro, Phe, Trp or Met"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Xaa Arg Xaa Ile Pro Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..4
- (D) OTHER INFORMATION: /note= "consensus telomerase RT sequence from motif 2"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 3
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid, Ala, Leu, Ile, Val, Pro, Phe, Trp or Met"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Phe Arg Xaa Ile  
1

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..6
- (D) OTHER INFORMATION: /note= "consensus telomerase RT sequence from motif A"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 2
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = charged amino acid, Asp, Glu, His, Lys or Arg"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 6
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid, Ala, Leu, Ile, Val, Pro, Phe, Trp or Met"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Pro Xaa Leu Tyr Phe Xaa  
1 5

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..6
- (D) OTHER INFORMATION: /note= "consensus telomerase RT sequence from motif B'"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gly Ile Pro Gln Gly Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..4
- (D) OTHER INFORMATION: /note= "consensus telomerase RT sequence from motif C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Leu Leu Arg Leu  
1

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..7  
 (D) OTHER INFORMATION: /note= "consensus telomerase RT sequence  
 from motif C"

(ix) FEATURE:  
 (A) NAME/KEY: Modified-site  
 (B) LOCATION: 5  
 (D) OTHER INFORMATION: /product= "OTHER"  
 /note= "Xaa = hydrophobic amino acid,  
 Ala, Leu, Ile, Val, Pro, Phe, Trp  
 or Met"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Asp Asp Phe Leu Xaa Ile Thr  
 1 5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 48 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
 (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..48  
 (D) OTHER INFORMATION: /note= "motif T peptide from  
 Schizosaccharomyces pombe TRT tezl p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Trp Leu Tyr Asn Ser Phe Ile Ile Pro Ile Leu Gln Ser Phe Phe Tyr  
 1 5 10 15

Ile Thr Glu Ser Ser Asp Leu Arg Asn Arg Thr Val Tyr Phe Arg Lys  
 20 25 30

Asp Ile Trp Lys Leu Leu Cys Arg Pro Phe Ile Thr Ser Met Lys Met  
 35 40 45

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 54 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..54
- (D) OTHER INFORMATION: /note= "motif 1 and 2 peptide from Schizosaccharomyces pombe TRT tezl1p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Asn Asn Val Arg Met Asp Thr Gln Lys Thr Thr Leu Pro Pro Ala Val  
1 5 10 15  
Ile Arg Leu Leu Pro Lys Lys Asn Thr Phe Arg Leu Ile Thr Asn Leu  
20 25 30  
Arg Lys Arg Phe Leu Ile Lys Met Gly Ser Asn Lys Lys Met Leu Val  
35 40 45  
Ser Thr Asn Gln Thr Leu  
50

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..34
- (D) OTHER INFORMATION: /note= "motif A peptide from Schizosaccharomyces pombe TRT tezl1p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Phe Gly Arg Lys Lys Tyr Phe Val Arg Ile Asp Ile Lys Ser Cys Tyr  
1 5 10 15  
Asp Arg Ile Lys Gln Asp Leu Met Phe Arg Ile Val Lys Lys Lys Leu  
20 25 30  
Lys Asp

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

[illegible]

Ser Gln Tyr Leu Gln Lys Val Gly Ile Pro Gln Gly Ser Ile Leu Ser  
1 5 10 15

Ser Phe Thr  
35

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 43 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(A) NAME/KEY: Peptide  
(B) LOCATION: 1..43  
(D) OTHER INFORMATION: /note= "motif C and D peptide from  
Schizosaccharomyces pombe TRT tezlp"

Leu Leu Arg Val Val Asp Asp Phe Leu Phe Ile Thr Val Asn Lys Lys  
1 5 10 15

Asn Phe Ser Thr Ser Leu Glu Lys Thr Val Ile  
35 40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

```
(A) NAME/KEY: Peptide
(B) LOCATION: 1..11
```



(D) OTHER INFORMATION: /note= "motif E peptide from  
Schizosaccharomyces pombe TRT tezl p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Lys Lys Arg Met Pro Phe Phe Gly Phe Ser Val  
1 5 10

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..48
- (D) OTHER INFORMATION: /note= "motif T peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe Tyr  
1 5 10 15  
Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys  
20 25 30  
Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu Lys  
35 40 45

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..54
- (D) OTHER INFORMATION: /note= "motif 1 and 2 peptide from  
human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg  
1 5 10 15

Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met  
20 25 30  
Asp Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu  
35 40 45  
Arg Leu Thr Ser Arg Val  
50

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:
  - (A) NAME/KEY: Peptide
  - (B) LOCATION: 1..34
  - (D) OTHER INFORMATION: /note= "motif A peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Pro Pro Pro Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr  
1 5 10 15  
Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile  
20 25 30  
Lys Pro

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:
  - (A) NAME/KEY: Peptide
  - (B) LOCATION: 1..35
  - (D) OTHER INFORMATION: /note= "motif B' peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln Gly Ser Ile Leu Ser  
1 5 10 15  
Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met Glu Asn Lys Leu Phe  
20 25 30

Ala Gly Ile  
35

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..43
- (D) OTHER INFORMATION: /note= "motif C and D peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Leu Leu Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr  
1                      5                      10                      15

His Ala Lys Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr  
20                      25                      30

Gly Cys Val Val Asn Leu Arg Lys Thr Val Val  
35                      40

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "motif E peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

His Gly Leu Phe Pro Trp Cys Gly Leu Leu Leu  
1                      5                      10

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid

(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide  
(B) LOCATION: 1..48  
(D) OTHER INFORMATION: /note= "motif T peptide from Euplotes  
aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Trp	Ile	Phe	Glu	Asp	Leu	Val	Val	Ser	Leu	Ile	Arg	Cys	Phe	Phe	Tyr
1				5					10					15	
Val	Thr	Glu	Gln	Gln	Lys	Ser	Tyr	Ser	Lys	Thr	Tyr	Tyr	Tyr	Arg	Lys
			20					25					30		
Asn	Ile	Trp	Asp	Val	Ile	Met	Lys	Met	Ser	Ile	Ala	Asp	Leu	Lys	Lys
			35				40					45			

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide  
(B) LOCATION: 1..54  
(D) OTHER INFORMATION: /note= "motif 1 and 2 peptide from  
Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Lys	Glu	Val	Glu	Glu	Trp	Lys	Lys	Ser	Leu	Gly	Phe	Ala	Pro	Gly	Lys
1				5					10					15	
Leu	Arg	Leu	Ile	Pro	Lys	Lys	Thr	Thr	Phe	Arg	Pro	Ile	Met	Thr	Phe
			20				25					30			
Asn	Lys	Lys	Ile	Val	Asn	Ser	Asp	Arg	Lys	Thr	Thr	Lys	Leu	Thr	Thr
		35				40					45				
Asn	Thr	Lys	Leu	Leu	Asn										
			50												

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 amino acids

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide  
(B) LOCATION: 1..34  
(D) OTHER INFORMATION: /note= "motif A peptide from Euplotes  
aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Gly	Gln	Pro	Lys	Leu	Phe	Phe	Ala	Thr	Met	Asp	Ile	Glu	Lys	Cys	Tyr
1			5					10					15		
Asp	Ser	Val	Asn	Arg	Glu	Lys	Leu	Ser	Thr	Phe	Leu	Lys	Thr	Thr	Lys
			20					25				30			
Leu	Leu														

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide  
(B) LOCATION: 1..35  
(D) OTHER INFORMATION: /note= "motif B' peptide from Euplotes  
aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Lys	Phe	Tyr	Lys	Gln	Thr	Lys	Gly	Ile	Pro	Gln	Gly	Leu	Cys	Val	Ser
1			5					10					15		
Ser	Ile	Leu	Ser	Ser	Phe	Tyr	Tyr	Ala	Thr	Leu	Glu	Glu	Ser	Ser	Leu
			20					25					30		
Gly	Phe	Leu													
			35												

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..43

(D) OTHER INFORMATION: /note= "motif C and D peptide from  
Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Leu Met Arg Leu Thr Asp Asp Tyr Leu Leu Ile Thr Thr Gln Glu Asn  
1 5 10 15

Asn Ala Val Leu Phe Ile Glu Lys Leu Ile Asn Val Ser Arg Glu Asn  
20 25 30

Gly Phe Lys Phe Asn Met Lys Lys Leu Gln Thr  
35 40

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..11

(D) OTHER INFORMATION: /note= "motif E peptide from Euplotes  
aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Gln Asp Tyr Cys Asp Trp Ile Gly Ile Ser Ile  
1 5 10

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..47

(D) OTHER INFORMATION: /note= "motif T peptide from  
Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Trp Leu Phe Arg Gln Leu Ile Pro Lys Ile Ile Gln Thr Phe Phe Tyr  
1 5 10 15  
Cys Thr Glu Ile Ser Ser Thr Val Thr Ile Val Tyr Phe Arg His Asp  
20 25 30  
Thr Trp Asn Lys Leu Ile Thr Pro Phe Ile Val Glu Tyr Phe Lys  
35 40 45

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..26  
(D) OTHER INFORMATION: /note= "motif 1 peptide from  
Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Cys Arg Asn His Asn Ser Tyr Thr Leu Ser Asn Phe Asn His Ser Lys  
1 5 10 15  
Met Arg Ile Ile Pro Lys Lys Ser Asn Asn  
20 25

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..29  
(D) OTHER INFORMATION: /note= "motif 2 peptide from  
Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Phe Arg Ile Ile Ala Ile Pro Cys Arg Gly Ala Asp Glu Glu Glu Phe  
1 5 10 15

Thr Ile Tyr Lys Glu Asn His Lys Asn Ala Ile Gln Pro  
 20 25

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 34 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..34  
 (D) OTHER INFORMATION: /note= "motif A peptide from  
 Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Val Leu Pro Glu Leu Tyr Phe Met Lys Phe Asp Val Lys Ser Cys Tyr  
 1 5 10 15  
 Asp Ser Ile Pro Arg Met Glu Cys Met Arg Ile Leu Lys Asp Ala Leu  
 20 25 30  
 Lys Asn

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..35  
 (D) OTHER INFORMATION: /note= "motif B' peptide from  
 Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Lys Cys Tyr Ile Arg Glu Asp Gly Leu Phe Gln Gly Ser Ser Leu Ser  
 1 5 10 15  
 Ala Pro Ile Val Asp Leu Val Tyr Asp Asp Leu Leu Glu Phe Tyr Ser  
 20 25 30  
 Glu Phe Lys  
 35



(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 43 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..43  
(D) OTHER INFORMATION: /note= "motif C and D peptide from  
Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Ile Leu Lys Leu Ala Asp Asp Phe Leu Ile Ile Ser Thr Asp Gln Gln  
1 5 10 15  
Gln Val Ile Asn Ile Lys Lys Leu Ala Met Gly Gly Phe Gln Lys Tyr  
20 25 30  
Asn Ala Lys Ala Asn Arg Asp Lys Ile Leu Ala  
35 40

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..11  
(D) OTHER INFORMATION: /note= "motif E peptide from  
Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Lys Glu Leu Glu Val Trp Lys His Ser Ser Thr  
1 5 10

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..4

(D) OTHER INFORMATION: /note= "consensus non-telomerase RT  
sequence from motif B'"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 1

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid,  
Ala, Leu, Ile, Val, Pro, Phe, Trp  
or Met"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Xaa Pro Gln Gly

1

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..5

(D) OTHER INFORMATION: /note= "consensus non-telomerase RT  
sequence from motif C"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 3

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid,  
Ala, Leu, Ile, Val, Pro, Phe, Trp  
or Met"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 4

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid,  
Ala, Leu, Ile, Val, Pro, Phe, Trp  
or Met"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 5

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid,

Ala, Leu, Ile, Val, Pro, Phe, Trp  
or Met"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Asp Asp Xaa Xaa Xaa  
1 5

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 55 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
(B) LOCATION: 1..55  
(D) OTHER INFORMATION: /note= "motif 1 and 2 peptide from  
Saccharomyces cerevisiae cytochrome  
oxidase group II intron 1-encoded  
mitochondrial protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Leu Ser Asn Glu Leu Gly Thr Gly Lys Phe Lys Phe Lys Pro Met Arg  
1 5 10 15  
Ile Val Asn Ile Pro Lys Pro Lys Gly Gly Ile Arg Pro Leu Ser Val  
20 25 30  
Gly Asn Pro Arg Asp Lys Ile Val Gln Glu Val Met Arg Met Ile Leu  
35 40 45  
Asp Thr Ile Phe Asp Lys Lys  
50 55

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
(B) LOCATION: 1..34  
(D) OTHER INFORMATION: /note= "motif A peptide from  
Saccharomyces cerevisiae cytochrome  
oxidase group II intron 1-encoded

mitochondrial protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Phe Gly Gly Ser Asn Trp Phe Ile Glu Val Asp Leu Lys Lys Cys Phe  
1                   5                   10                   15  
Asp Thr Ile Ser His Asp Leu Ile Ile Lys Glu Leu Lys Arg Tyr Ile  
          20                   25                   30  
Ser Asp

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..35
- (D) OTHER INFORMATION: /note= "motif B' peptide from  
Saccharomyces cerevisiae cytochrome  
oxidase group II intron 1-encoded  
mitochondrial protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Thr Tyr His Lys Pro Met Leu Gly Leu Pro Gln Gly Ser Leu Ile Ser  
1                   5                   10                   15  
Pro Ile Leu Cys Asn Ile Val Met Thr Leu Val Asp Asn Trp Leu Glu  
          20                   25                   30  
Asp Tyr Ile  
          35

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..17
- (D) OTHER INFORMATION: /note= "motif C peptide from  
Saccharomyces cerevisiae cytochrome

oxidase group II intron 1-encoded  
mitochondrial protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Tyr Val Arg Tyr Ala Asp Asp Ile Leu Ile Gly Val Leu Gly Ser Lys  
1                   5                   10                   15

Asn

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..25
- (D) OTHER INFORMATION: /note= "motif D peptide from  
Saccharomyces cerevisiae cytochrome  
oxidase group II intron 1-encoded  
mitochondrial protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Lys Met Ile Lys Arg Asp Leu Asn Asn Phe Leu Asn Ser Leu Gly Leu  
1                   5                   10                   15

Thr Ile Asn Glu Glu Lys Thr Leu Ile  
20                   25

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "motif E peptide from  
Saccharomyces cerevisiae cytochrome  
oxidase group II intron 1-encoded  
mitochondrial protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Glu Thr Pro Ala Arg Phe Leu Gly Tyr Asn Ile  
 1 5 10

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..26
- (D) OTHER INFORMATION: /note= "motif 1 peptide from Drosophila melanogaster TART non-LTR retrotransposable element reverse transcriptase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Ser Ile Leu Arg Ile Gly Tyr Tyr Pro Asp Ala Trp Lys His Ala Gln  
 1 5 10 15

Val Lys Met Ile Leu Lys Pro Gly Lys Ser  
 20 25

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..29
- (D) OTHER INFORMATION: /note= "motif 2 peptide from Drosophila melanogaster TART non-LTR retrotransposable element reverse transcriptase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Tyr Arg Pro Ile Ser Leu Leu Ser Gly Leu Ser Lys Met Phe Glu Arg  
 1 5 10 15

Leu Leu Leu Lys Arg Leu Phe Arg Val Asp Leu Phe Lys  
 20 25

(D) OTHER INFORMATION: /note= "motif 2 and A peptide from HIV-1 reverse transcriptase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln	Asp
1				5					10					15	
Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys	Lys
			20					25					30		
Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser	Val
		35					40					45			
Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro	
	50					55					60				

(2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
(B) LOCATION: 35  
(D) OTHER INFORMATION: /note= "motif B' peptide from HIV-1 reverse transcriptase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser
1				5					10					15	
Pro	Ala	Ile	Phe	Gln	Ser	Ser	Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Lys
			20					25					30		
Lys	Gln	Asn													
		35													

(2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide

(B) LOCATION: 1..17  
(D) OTHER INFORMATION: /note= "motif C peptide from HIV-1 reverse transcriptase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile  
1                      5                      10                      15

Gly

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 37 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..37  
(D) OTHER INFORMATION: /note= "motif D and E peptide from HIV-1 reverse transcriptase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly  
1                      5                      10                      15

Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp  
20                      25                      30

Met Gly Ile Thr Leu  
35

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..4  
(D) OTHER INFORMATION: /note= "consensus telomerase RT finger sequence from motif 1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:



Ile Pro Lys Lys

1

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..5
- (D) OTHER INFORMATION: /note= "consensus telomerase RT palm,  
primer grip sequence from motif C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Leu Leu Leu Arg Leu

1

5

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..4
- (D) OTHER INFORMATION: /note= "consensus telomerase RT palm,  
primer grip sequence from motif C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

Asp Asp Phe Leu

1

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..43
- (D) OTHER INFORMATION: /note= "telomerase specific motif T peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe Tyr  
1                   5                   10                   15  
Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys  
                  20                   25                   30  
Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile  
          35                   40

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..5
- (D) OTHER INFORMATION: /note= "telomerase specific motif T' peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Glu Ala Glu Val Arg  
1                   5

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..16
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif 1 and 2 peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val  
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:75:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..17
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif A peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Pro Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr  
 1 5 10 15  
 Ile

(2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif B' peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Tyr Val Gln Cys Gln Gly Ile Pro Gln Gly Ser Ile Leu Ser Thr Leu  
 1 5 10 15  
 Leu Cys Ser Leu Cys Tyr  
 20

(2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 13 amino acids
  - (B) TYPE: amino acid

- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..13
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip motif C peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Leu Leu Leu Arg Leu Val Asp Asp Phe Leu Leu Val Thr  
1 5 10

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..16
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip motif D peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Gly Val Pro Glu Tyr Gly Cys Val Val Asn Leu Arg Lys Thr Val Val  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip motif E peptide from human TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Trp Cys Gly Leu Leu Asp Thr Arg Thr Leu  
 1 5 10

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..43
- (D) OTHER INFORMATION: /note= "telomerase specific motif T peptide from Schizosaccharomyces pombe TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Trp Leu Tyr Asn Ser Phe Ile Ile Pro Ile Leu Gln Ser Phe Phe Tyr  
 1 5 10 15  
 Ile Thr Glu Ser Ser Asp Leu Arg Asn Arg Thr Val Tyr Phe Arg Lys  
 20 25 30  
 Asp Ile Trp Lys Leu Leu Cys Arg Pro Phe Ile  
 35 40

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..5
- (D) OTHER INFORMATION: /note= "telomerase specific motif T' peptide from Schizosaccharomyces pombe TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Glu Asn Asn Val Arg  
 1 5

(2) INFORMATION FOR SEQ ID NO:82:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 16 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
 (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..16  
 (D) OTHER INFORMATION: /note= "telomerase RT finger motif  
 1 and 2 peptide from Schizosaccharomyces  
 pombe TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Ala	Val	Ile	Arg	Leu	Leu	Pro	Lys	Lys	Asn	Thr	Phe	Arg	Leu	Ile	Thr
1				5					10					15	

(2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 17 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
 (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..17  
 (D) OTHER INFORMATION: /note= "telomerase RT finger motif A  
 peptide from Schizosaccharomyces  
 pombe TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Arg	Lys	Lys	Tyr	Phe	Val	Arg	Ile	Asp	Ile	Lys	Ser	Cys	Tyr	Asp	Arg
1				5					10					15	

Ile

(2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif B'  
peptide from Schizosaccharomyces  
pombe TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Tyr Leu Gln Lys Val Gly Ile Pro Gln Gly Ser Ile Leu Ser Ser Phe  
1                      5                      10                      15  
Leu Cys His Phe Tyr Met  
                    20

(2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 13 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:
  - (A) NAME/KEY: Peptide
  - (B) LOCATION: 1..13
  - (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip  
motif C peptide from Schizosaccharomyces  
pombe TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Val Leu Leu Arg Val Val Asp Asp Phe Leu Phe Ile Thr  
1                      5                      10

(2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:
  - (A) NAME/KEY: Peptide
  - (B) LOCATION: 1..16
  - (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip  
motif D peptide from Schizosaccharomyces  
pombe TRT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Gly Phe Glu Lys His Asn Phe Ser Thr Ser Leu Glu Lys Thr Val Ile

1

5

10

15

## (2) INFORMATION FOR SEQ ID NO:87:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip motif E peptide from Schizosaccharomyces pombe TRT"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Phe Phe Gly Phe Ser Val Asn Met Arg Ser Leu  
 1                      5                      10

## (2) INFORMATION FOR SEQ ID NO:88:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..43
- (D) OTHER INFORMATION: /note= "telomerase specific motif T peptide from Euplotes aediculatus p123"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Trp Ile Phe Glu Asp Leu Val Val Ser Leu Ile Arg Cys Phe Phe Tyr  
 1                      5                      10                      15

Val Thr Glu Gln Gln Lys Ser Tyr Ser Lys Thr Tyr Tyr Tyr Arg Lys  
                     20                      25                      30

Asn Ile Trp Asp Val Ile Met Lys Met Ser Ile  
                     35                      40

## (2) INFORMATION FOR SEQ ID NO:89:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids



- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..5
- (D) OTHER INFORMATION: /note= "telomerase specific motif T'  
peptide from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Glu Lys Glu Val Glu  
1 5

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..16
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif  
1 and 2 peptide from Euplotes  
aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Gly Lys Leu Arg Leu Ile Pro Lys Lys Thr Phe Arg Pro Ile Met  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..17
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif A  
peptide from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Pro Lys Leu Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr Asp Ser  
1                      5                      10                      15  
Val

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif B'  
peptide from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

Tyr Lys Gln Thr Lys Gly Ile Pro Gln Gly Leu Cys Val Ser Ser Ile  
1                      5                      10                      15  
Leu Ser Ser Phe Tyr Tyr  
20

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..13
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip  
motif C peptide from Euplotes  
aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu Leu Ile Thr  
1                      5                      10

(2) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 16 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
 (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..16  
 (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip  
 motif D peptide from Euplotes  
 aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

Val	Ser	Arg	Glu	Asn	Gly	Phe	Lys	Phe	Asn	Met	Lys	Lys	Leu	Gln	Thr
1				5					10					15	

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
 (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..11  
 (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip  
 motif E peptide from Euplotes  
 aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

Trp	Ile	Gly	Ile	Ser	Ile	Asp	Met	Lys	Thr	Leu
1				5					10	

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
 (A) NAME/KEY: Peptide  
 (B) LOCATION: 1..42

(D) OTHER INFORMATION: /note= "telomerase specific motif T  
peptide from Saccharomyces  
cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

Trp	Leu	Phe	Arg	Gln	Leu	Ile	Pro	Lys	Ile	Ile	Gln	Thr	Phe	Phe	Tyr
1				5					10					15	
Cys	Thr	Glu	Ile	Ser	Ser	Thr	Val	Thr	Ile	Val	Tyr	Phe	Arg	His	Asp
		20						25					30		
Thr	Trp	Asn	Lys	Leu	Ile	Thr	Pro	Phe	Ile						
		35					40								

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..5
- (D) OTHER INFORMATION: /note= "telomerase specific motif T'  
peptide from Saccharomyces  
cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

Glu	Asn	Asn	Val	Cys
1				5

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif 1  
peptide from Saccharomyces  
cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Ser Lys Met Arg Ile Ile Pro Lys Lys Ser Asn  
1 5 10

(2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..5  
(D) OTHER INFORMATION: /note= "telomerase RT finger motif 2  
peptide from Saccharomyces  
cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

Phe Arg Ile Ile Ala  
1 5

(2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..17  
(D) OTHER INFORMATION: /note= "telomerase RT finger motif A  
peptide from Saccharomyces  
cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

Pro Glu Leu Tyr Phe Met Lys Phe Asp Val Lys Ser Cys Tyr Asp Ser  
1 5 10 15  
Ile

(2) INFORMATION FOR SEQ ID NO:101:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "telomerase RT finger motif B' peptide from Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

Tyr	Ile	Arg	Glu	Asp	Gly	Leu	Phe	Gln	Gly	Ser	Ser	Leu	Ser	Ala	Pro
1				5					10					15	
Ile	Val	Asp	Leu	Val	Tyr										
						20									

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..13
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip motif C peptide from Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

Leu	Ile	Leu	Lys	Leu	Ala	Asp	Asp	Phe	Leu	Ile	Ile	Ser
1				5					10			

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..16
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip"

motif D peptide from *Saccharomyces cerevisiae* EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

Gly Phe Gln Lys Tyr Asn Ala Lys Ala Asn Arg Asp Lys Ile Leu Ala  
1                   5                   10                   15

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "telomerase RT palm, primer grip motif E peptide from *Saccharomyces cerevisiae* EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

Trp Lys His Ser Ser Thr Met Asn Asn Phe His  
1                   5                   10

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: protein\_bind
- (B) LOCATION: 1..10
- (D) OTHER INFORMATION: /note= "NFkappaB CS1 binding site motif"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GGGRHTYYHC

10

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: protein\_bind
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "NFkappaB MHC I.2 binding site motif"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

TGGGCTTCCC C

11

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: protein\_bind
- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "NFkappaB CS2 binding site motif"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

RGGGRMTYYC C

11

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: protein\_bind
- (B) LOCATION: 1..18
- (D) OTHER INFORMATION: /note= "topoisomerase II cleavage site motif"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

RNYNNCNGY NGKTNYNY

18



(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3279 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 101..3196
- (D) OTHER INFORMATION: /note= "Euplotes aediculatus 123 kDa telomerase protein subunit (TRT)"  
/codon= (seq: "tga", aa: Cys)  
/product= "Euplotes aediculatus 123 kDa telomerase protein subunit (TRT)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

AAAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT AACCTCAGTA	60
TTAATAAGCT CAGATTTTAA ATATTAATTA CAAACCTAA ATG GAG GTT GAT GTT	115
Met Glu Val Asp Val	
1 5	
GAT AAT CAA GCT GAT AAT CAT GGC ATT CAC TCA GCT CTT AAG ACT TGT	163
Asp Asn Gln Ala Asp Asn His Gly Ile His Ser Ala Leu Lys Thr Cys	
10 15 20	
GAA GAA ATT AAA GAA GCT AAA ACG TTG TAC TCT TGG ATC CAG AAA GTT	211
Glu Glu Ile Lys Glu Ala Lys Thr Leu Tyr Ser Trp Ile Gln Lys Val	
25 30 35	
ATT AGA TGA AGA AAT CAA TCT CAA AGT CAT TAT AAA GAT TTA GAA GAT	259
Ile Arg Cys Arg Asn Gln Ser Gln Ser His Tyr Lys Asp Leu Glu Asp	
40 45 50	
ATT AAA ATA TTT GCG CAG ACA AAT ATT GTT GCT ACT CCA CGA GAC TAT	307
Ile Lys Ile Phe Ala Gln Thr Asn Ile Val Ala Thr Pro Arg Asp Tyr	
55 60 65	
AAT GAA GAA GAT TTT AAA GTT ATT GCA AGA AAA GAA GTA TTT TCA ACT	355
Asn Glu Glu Asp Phe Lys Val Ile Ala Arg Lys Glu Val Phe Ser Thr	
70 75 80 85	
GGA CTA ATG ATC GAA CTT ATT GAC AAA TGC TTA GTT GAA CTT CTT TCA	403
Gly Leu Met Ile Glu Leu Ile Asp Lys Cys Leu Val Glu Leu Leu Ser	
90 95 100	
TCA AGC GAT GTT TCA GAT AGA CAA AAA CTT CAA TGA TTT GGA TTT CAA	451
Ser Ser Asp Val Ser Asp Arg Gln Lys Leu Gln Cys Phe Gly Phe Gln	
105 110 115	
CTT AAG GGA AAT CAA TTA GCA AAG ACC CAT TTA TTA ACA GCT CTT TCA	499
Leu Lys Gly Asn Gln Leu Ala Lys Thr His Leu Leu Thr Ala Leu Ser	
120 125 130	
ACT CAA AAG CAG TAT TTC TTT CAA GAC GAA TGG AAC CAA GTT AGA GCA	547

Thr	Gln	Lys	Gln	Tyr	Phe	Phe	Gln	Asp	Glu	Trp	Asn	Gln	Val	Arg	Ala	
135						140					145					
ATG	ATT	GGA	AAT	GAG	CTC	TTC	CGA	CAT	CTC	TAC	ACT	AAA	TAT	TTA	ATA	595
Met	Ile	Gly	Asn	Glu	Leu	Phe	Arg	His	Leu	Tyr	Thr	Lys	Tyr	Leu	Ile	
150					155					160					165	
TTC	CAG	CGA	ACT	TCT	GAA	GGA	ACT	CTT	GTT	CAA	TTT	TGC	GGG	AAT	AAC	643
Phe	Gln	Arg	Thr	Ser	Glu	Gly	Thr	Leu	Val	Gln	Phe	Cys	Gly	Asn	Asn	
				170					175					180		
GTT	TTT	GAT	CAT	TTG	AAA	GTC	AAC	GAT	AAG	TTT	GAC	AAA	AAG	CAA	AAA	691
Val	Phe	Asp	His	Leu	Lys	Val	Asn	Asp	Lys	Phe	Asp	Lys	Lys	Gln	Lys	
			185					190					195			
GGT	GGA	GCA	GCA	GAC	ATG	AAT	GAA	CCT	CGA	TGT	TGA	TCA	ACC	TGC	AAA	739
Gly	Gly	Ala	Ala	Asp	Met	Asn	Glu	Pro	Arg	Cys	Cys	Ser	Thr	Cys	Lys	
		200					205					210				
TAC	AAT	GTC	AAG	AAT	GAG	AAA	GAT	CAC	TTT	CTC	AAC	AAC	ATC	AAC	GTG	787
Tyr	Asn	Val	Lys	Asn	Glu	Lys	Asp	His	Phe	Leu	Asn	Asn	Ile	Asn	Val	
	215					220					225					
CCG	AAT	TGG	AAT	AAT	ATG	AAA	TCA	AGA	ACC	AGA	ATA	TTT	TAT	TGC	ACT	835
Pro	Asn	Trp	Asn	Asn	Met	Lys	Ser	Arg	Thr	Arg	Ile	Phe	Tyr	Cys	Thr	
230					235					240					245	
CAT	TTT	AAT	AGA	AAT	AAC	CAA	TTC	TTC	AAA	AAG	CAT	GAG	TTT	GTG	AGT	883
His	Phe	Asn	Arg	Asn	Asn	Gln	Phe	Phe	Lys	Lys	His	Glu	Phe	Val	Ser	
				250					255					260		
AAC	AAA	AAC	AAT	ATT	TCA	GCG	ATG	GAC	AGA	GCT	CAG	ACG	ATA	TTC	ACG	931
Asn	Lys	Asn	Asn	Ile	Ser	Ala	Met	Asp	Arg	Ala	Gln	Thr	Ile	Phe	Thr	
			265					270					275			
AAT	ATA	TTC	AGA	TTT	AAT	AGA	ATT	AGA	AAG	AAG	CTA	AAA	GAT	AAG	GTT	979
Asn	Ile	Phe	Arg	Phe	Asn	Arg	Ile	Arg	Lys	Lys	Leu	Lys	Asp	Lys	Val	
		280					285					290				
ATC	GAA	AAA	ATT	GCC	TAC	ATG	CTT	GAG	AAA	GTC	AAA	GAT	TTT	AAC	TTC	1027
Ile	Glu	Lys	Ile	Ala	Tyr	Met	Leu	Glu	Lys	Val	Lys	Asp	Phe	Asn	Phe	
	295					300					305					
AAC	TAC	TAT	TTA	ACA	AAA	TCT	TGT	CCT	CTT	CCA	GAA	AAT	TGG	CGG	GAA	1075
Asn	Tyr	Tyr	Leu	Thr	Lys	Ser	Cys	Pro	Leu	Pro	Glu	Asn	Trp	Arg	Glu	
310					315					320					325	
CGG	AAA	CAA	AAA	ATC	GAA	AAC	TTG	ATA	AAT	AAA	ACT	AGA	GAA	GAA	AAG	1123
Arg	Lys	Gln	Lys	Ile	Glu	Asn	Leu	Ile	Asn	Lys	Thr	Arg	Glu	Glu	Lys	
				330					335					340		
TCG	AAG	TAC	TAT	GAA	GAG	CTG	TTT	AGC	TAC	ACA	ACT	GAT	AAT	AAA	TGC	1171
Ser	Lys	Tyr	Tyr	Glu	Glu	Leu	Phe	Ser	Tyr	Thr	Thr	Asp	Asn	Lys	Cys	
			345					350					355			
GTC	ACA	CAA	TTT	ATT	AAT	GAA	TTT	TTC	TAC	AAT	ATA	CTC	CCC	AAA	GAC	1219
Val	Thr	Gln	Phe	Ile	Asn	Glu	Phe	Phe	Tyr	Asn	Ile	Leu	Pro	Lys	Asp	
		360					365					370				
TTT	TTG	ACT	GGA	AGA	AAC	CGT	AAG	AAT	TTT	CAA	AAG	AAA	GTT	AAG	AAA	1267
Phe	Leu	Thr	Gly	Arg	Asn	Arg	Lys	Asn	Phe	Gln	Lys	Lys	Val	Lys	Lys	
	375					380					385					

TAT GTG GAA CTA AAC AAG CAT GAA CTC ATT CAC AAA AAC TTA TTG CTT	1315
Tyr Val Glu Leu Asn Lys His Glu Leu Ile His Lys Asn Leu Leu Leu	
390 395 400 405	
GAG AAG ATC AAT ACA AGA GAA ATA TCA TGG ATG CAG GTT GAG ACC TCT	1363
Glu Lys Ile Asn Thr Arg Glu Ile Ser Trp Met Gln Val Glu Thr Ser	
410 415 420	
GCA AAG CAT TTT TAT TAT TTT GAT CAC GAA AAC ATC TAC GTC TTA TGG	1411
Ala Lys His Phe Tyr Tyr Phe Asp His Glu Asn Ile Tyr Val Leu Trp	
425 430 435	
AAA TTG CTC CGA TGG ATA TTC GAG GAT CTC GTC GTC TCG CTG ATT AGA	1459
Lys Leu Leu Arg Trp Ile Phe Glu Asp Leu Val Val Ser Leu Ile Arg	
440 445 450	
TGA TTT TTC TAT GTC ACC GAG CAA CAG AAA AGT TAC TCC AAA ACC TAT	1507
Cys Phe Phe Tyr Val Thr Glu Gln Gln Lys Ser Tyr Ser Lys Thr Tyr	
455 460 465	
TAC TAC AGA AAG AAT ATT TGG GAC GTC ATT ATG AAA ATG TCA ATC GCA	1555
Tyr Tyr Arg Lys Asn Ile Trp Asp Val Ile Met Lys Met Ser Ile Ala	
470 475 480 485	
GAC TTA AAG AAG GAA ACG CTT GCT GAG GTC CAA GAA AAA GAG GTT GAA	1603
Asp Leu Lys Lys Glu Thr Leu Ala Glu Val Gln Glu Lys Glu Val Glu	
490 495 500	
GAA TGG AAA AAG TCG CTT GGA TTT GCA CCT GGA AAA CTC AGA CTA ATA	1651
Glu Trp Lys Lys Ser Leu Gly Phe Ala Pro Gly Lys Leu Arg Leu Ile	
505 510 515	
CCG AAG AAA ACT ACT TTC CGT CCA ATT ATG ACT TTC AAT AAG AAG ATT	1699
Pro Lys Lys Thr Thr Phe Arg Pro Ile Met Thr Phe Asn Lys Lys Ile	
520 525 530	
GTA AAT TCA GAC CGG AAG ACT ACA AAA TTA ACT ACA AAT ACG AAG TTA	1747
Val Asn Ser Asp Arg Lys Thr Thr Lys Leu Thr Thr Asn Thr Lys Leu	
535 540 545	
TTG AAC TCT CAC TTA ATG CTT AAG ACA TTG AAG AAT AGA ATG TTT AAA	1795
Leu Asn Ser His Leu Met Leu Lys Thr Leu Lys Asn Arg Met Phe Lys	
550 555 560 565	
GAT CCT TTT GGA TTC GCT GTT TTT AAC TAT GAT GAT GTA ATG AAA AAG	1843
Asp Pro Phe Gly Phe Ala Val Phe Asn Tyr Asp Asp Val Met Lys Lys	
570 575 580	
TAT GAG GAG TTT GTT TGC AAA TGG AAG CAA GTT GGA CAA CCA AAA CTC	1891
Tyr Glu Glu Phe Val Cys Lys Trp Lys Gln Val Gly Gln Pro Lys Leu	
585 590 595	
TTC TTT GCA ACT ATG GAT ATC GAA AAG TGA TAT GAT AGT GTA AAC AGA	1939
Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr Asp Ser Val Asn Arg	
600 605 610	
GAA AAA CTA TCA ACA TTC CTA AAA ACT ACT AAA TTA CTT TCT TCA GAT	1987
Glu Lys Leu Ser Thr Phe Leu Lys Thr Thr Lys Leu Leu Ser Ser Asp	
615 620 625	
TTC TGG ATT ATG ACT GCA CAA ATT CTA AAG AGA AAG AAT AAC ATA GTT	2035
Phe Trp Ile Met Thr Ala Gln Ile Leu Lys Arg Lys Asn Asn Ile Val	

630		635		640		645	
ATC GAT TCG AAA AAC TTT AGA AAG AAA GAA ATG AAA GAT TAT TTT AGA	2083						
Ile Asp Ser Lys Asn Phe Arg Lys Lys Glu Met Lys Asp Tyr Phe Arg							
		650		655		660	
CAG AAA TTC CAG AAG ATT GCA CTT GAA GGA GGA CAA TAT CCA ACC TTA	2131						
Gln Lys Phe Gln Lys Ile Ala Leu Glu Gly Gly Gln Tyr Pro Thr Leu				670		675	
TTC AGT GTT CTT GAA AAT GAA CAA AAT GAC TTA AAT GCA AAG AAA ACA	2179						
Phe Ser Val Leu Glu Asn Glu Gln Asn Asp Leu Asn Ala Lys Lys Thr				685		690	
TTA ATT GTT GAA GCA AAG CAA AGA AAT TAT TTT AAG AAA GAT AAC TTA	2227						
Leu Ile Val Glu Ala Lys Gln Arg Asn Tyr Phe Lys Lys Asp Asn Leu				700		705	
CTT CAA CCA GTC ATT AAT ATT TGC CAA TAT AAT TAC ATT AAC TTT AAT	2275						
Leu Gln Pro Val Ile Asn Ile Cys Gln Tyr Asn Tyr Ile Asn Phe Asn				710		715	
						720	725
GGG AAG TTT TAT AAA CAA ACA AAA GGA ATT CCT CAA GGT CTT TGA GTT	2323						
Gly Lys Phe Tyr Lys Gln Thr Lys Gly Ile Pro Gln Gly Leu Cys Val						730	735
							740
TCA TCA ATT TTG TCA TCA TTT TAT TAT GCA ACA TTA GAG GAA AGC TCC	2371						
Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala Thr Leu Glu Glu Ser Ser						745	750
							755
TTA GGA TTC CTT AGA GAT GAA TCA ATG AAC CCT GAA AAT CCA AAT GTT	2419						
Leu Gly Phe Leu Arg Asp Glu Ser Met Asn Pro Glu Asn Pro Asn Val						760	765
							770
AAT CTT CTA ATG AGA CTT ACA GAT GAC TAT CTT TTG ATT ACA ACT CAA	2467						
Asn Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu Leu Ile Thr Thr Gln						775	780
							785
GAG AAT AAT GCA GTA TTG TTT ATT GAG AAA CTT ATA AAC GTA AGT CGT	2515						
Glu Asn Asn Ala Val Leu Phe Ile Glu Lys Leu Ile Asn Val Ser Arg						790	795
							800
GAA AAT GGA TTT AAA TTC AAT ATG AAG AAA CTA CAG ACT AGT TTT CCA	2563						
Glu Asn Gly Phe Lys Phe Asn Met Lys Lys Leu Gln Thr Ser Phe Pro						810	815
							820
TTA AGT CCA AGC AAA TTT GCA AAA TAC GGA ATG GAT AGT GTT GAG GAG	2611						
Leu Ser Pro Ser Lys Phe Ala Lys Tyr Gly Met Asp Ser Val Glu Glu						825	830
							835
CAA AAT ATT GTT CAA GAT TAC TGC GAT TGG ATT GGC ATC TCA ATT GAT	2659						
Gln Asn Ile Val Gln Asp Tyr Cys Asp Trp Ile Gly Ile Ser Ile Asp						840	845
							850
ATG AAA ACT CTT GCT TTA ATG CCA AAT ATT AAC TTG AGA ATA GAA GGA	2707						
Met Lys Thr Leu Ala Leu Met Pro Asn Ile Asn Leu Arg Ile Glu Gly						855	860
							865
ATT CTG TGT ACA CTC AAT CTA AAC ATG CAA ACA AAG AAA GCA TCA ATG	2755						
Ile Leu Cys Thr Leu Asn Leu Asn Met Gln Thr Lys Lys Ala Ser Met						870	875
							880
							885

TGG CTC AAG AAG AAA CTA AAG TCG TTT TTA ATG AAT AAC ATT ACC CAT	2803
Trp Leu Lys Lys Lys Leu Lys Ser Phe Leu Met Asn Asn Ile Thr His	
890 895 900	
TAT TTT AGA AAG ACG ATT ACA ACC GAA GAC TTT GCG AAT AAA ACT CTC	2851
Tyr Phe Arg Lys Thr Ile Thr Thr Glu Asp Phe Ala Asn Lys Thr Leu	
905 910 915	
AAC AAG TTA TTT ATA TCA GGC GGT TAC AAA TAC ATG CAA TGA GCC AAA	2899
Asn Lys Leu Phe Ile Ser Gly Gly Tyr Lys Tyr Met Gln Cys Ala Lys	
920 925 930	
GAA TAC AAG GAC CAC TTT AAG AAG AAC TTA GCT ATG AGC AGT ATG ATC	2947
Glu Tyr Lys Asp His Phe Lys Lys Asn Leu Ala Met Ser Ser Met Ile	
935 940 945	
GAC TTA GAG GTA TCT AAA ATT ATA TAC TCT GTA ACC AGA GCA TTC TTT	2995
Asp Leu Glu Val Ser Lys Ile Ile Tyr Ser Val Thr Arg Ala Phe Phe	
950 955 960 965	
AAA TAC CTT GTG TGC AAT ATT AAG GAT ACA ATT TTT GGA GAG GAG CAT	3043
Lys Tyr Leu Val Cys Asn Ile Lys Asp Thr Ile Phe Gly Glu Glu His	
970 975 980	
TAT CCA GAC TTT TTC CTT AGC ACA CTG AAG CAC TTT ATT GAA ATA TTC	3091
Tyr Pro Asp Phe Phe Leu Ser Thr Leu Lys His Phe Ile Glu Ile Phe	
985 990 995	
AGC ACA AAA AAG TAC ATT TTC AAC AGA GTT TGC ATG ATC CTC AAG GCA	3139
Ser Thr Lys Lys Tyr Ile Phe Asn Arg Val Cys Met Ile Leu Lys Ala	
1000 1005 1010	
AAA GAA GCA AAG CTA AAA AGT GAC CAA TGT CAA TCT CTA ATT CAA TAT	3187
Lys Glu Ala Lys Leu Lys Ser Asp Gln Cys Gln Ser Leu Ile Gln Tyr	
1015 1020 1025	
GAT GCA TAGTCGACTA TTCTAACTTA TTTTGGAAAG TTAATTTTCA ATTTTGTCT	3243
Asp Ala	
1030	
TATATACTGG GGTTTTGGGG TTTTGGGGTT TTGGGG	3279

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1031 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

Met	Glu	Val	Asp	Val	Asp	Asn	Gln	Ala	Asp	Asn	His	Gly	Ile	His	Ser
1				5					10					15	
Ala	Leu	Lys	Thr	Cys	Glu	Glu	Ile	Lys	Glu	Ala	Lys	Thr	Leu	Tyr	Ser
			20					25					30		
Trp	Ile	Gln	Lys	Val	Ile	Arg	Cys	Arg	Asn	Gln	Ser	Gln	Ser	His	Tyr
		35					40					45			

Lys Asp Leu Glu Asp Ile Lys Ile Phe Ala Gln Thr Asn Ile Val Ala  
 50 55 60  
 Thr Pro Arg Asp Tyr Asn Glu Glu Asp Phe Lys Val Ile Ala Arg Lys  
 65 70 75 80  
 Glu Val Phe Ser Thr Gly Leu Met Ile Glu Leu Ile Asp Lys Cys Leu  
 85 90 95  
 Val Glu Leu Leu Ser Ser Ser Asp Val Ser Asp Arg Gln Lys Leu Gln  
 100 105 110  
 Cys Phe Gly Phe Gln Leu Lys Gly Asn Gln Leu Ala Lys Thr His Leu  
 115 120 125  
 Leu Thr Ala Leu Ser Thr Gln Lys Gln Tyr Phe Phe Gln Asp Glu Trp  
 130 135 140  
 Asn Gln Val Arg Ala Met Ile Gly Asn Glu Leu Phe Arg His Leu Tyr  
 145 150 155 160  
 Thr Lys Tyr Leu Ile Phe Gln Arg Thr Ser Glu Gly Thr Leu Val Gln  
 165 170 175  
 Phe Cys Gly Asn Asn Val Phe Asp His Leu Lys Val Asn Asp Lys Phe  
 180 185 190  
 Asp Lys Lys Gln Lys Gly Gly Ala Ala Asp Met Asn Glu Pro Arg Cys  
 195 200 205  
 Cys Ser Thr Cys Lys Tyr Asn Val Lys Asn Glu Lys Asp His Phe Leu  
 210 215 220  
 Asn Asn Ile Asn Val Pro Asn Trp Asn Asn Met Lys Ser Arg Thr Arg  
 225 230 235 240  
 Ile Phe Tyr Cys Thr His Phe Asn Arg Asn Asn Gln Phe Phe Lys Lys  
 245 250 255  
 His Glu Phe Val Ser Asn Lys Asn Asn Ile Ser Ala Met Asp Arg Ala  
 260 265 270  
 Gln Thr Ile Phe Thr Asn Ile Phe Arg Phe Asn Arg Ile Arg Lys Lys  
 275 280 285  
 Leu Lys Asp Lys Val Ile Glu Lys Ile Ala Tyr Met Leu Glu Lys Val  
 290 295 300  
 Lys Asp Phe Asn Phe Asn Tyr Tyr Leu Thr Lys Ser Cys Pro Leu Pro  
 305 310 315 320  
 Glu Asn Trp Arg Glu Arg Lys Gln Lys Ile Glu Asn Leu Ile Asn Lys  
 325 330 335  
 Thr Arg Glu Glu Lys Ser Lys Tyr Tyr Glu Glu Leu Phe Ser Tyr Thr  
 340 345 350  
 Thr Asp Asn Lys Cys Val Thr Gln Phe Ile Asn Glu Phe Phe Tyr Asn  
 355 360 365  
 Ile Leu Pro Lys Asp Phe Leu Thr Gly Arg Asn Arg Lys Asn Phe Gln  
 370 375 380

Lys	Lys	Val	Lys	Lys	Tyr	Val	Glu	Leu	Asn	Lys	His	Glu	Leu	Ile	His
385					390					395					400
Lys	Asn	Leu	Leu	Leu	Glu	Lys	Ile	Asn	Thr	Arg	Glu	Ile	Ser	Trp	Met
			405						410					415	
Gln	Val	Glu	Thr	Ser	Ala	Lys	His	Phe	Tyr	Tyr	Phe	Asp	His	Glu	Asn
			420					425					430		
Ile	Tyr	Val	Leu	Trp	Lys	Leu	Leu	Arg	Trp	Ile	Phe	Glu	Asp	Leu	Val
		435					440					445			
Val	Ser	Leu	Ile	Arg	Cys	Phe	Phe	Tyr	Val	Thr	Glu	Gln	Gln	Lys	Ser
	450					455					460				
Tyr	Ser	Lys	Thr	Tyr	Tyr	Tyr	Arg	Lys	Asn	Ile	Trp	Asp	Val	Ile	Met
465					470					475					480
Lys	Met	Ser	Ile	Ala	Asp	Leu	Lys	Lys	Glu	Thr	Leu	Ala	Glu	Val	Gln
				485					490						495
Glu	Lys	Glu	Val	Glu	Glu	Trp	Lys	Lys	Ser	Leu	Gly	Phe	Ala	Pro	Gly
			500					505					510		
Lys	Leu	Arg	Leu	Ile	Pro	Lys	Lys	Thr	Thr	Phe	Arg	Pro	Ile	Met	Thr
		515					520					525			
Phe	Asn	Lys	Lys	Ile	Val	Asn	Ser	Asp	Arg	Lys	Thr	Thr	Lys	Leu	Thr
	530					535					540				
Thr	Asn	Thr	Lys	Leu	Leu	Asn	Ser	His	Leu	Met	Leu	Lys	Thr	Leu	Lys
545					550					555					560
Asn	Arg	Met	Phe	Lys	Asp	Pro	Phe	Gly	Phe	Ala	Val	Phe	Asn	Tyr	Asp
				565					570					575	
Asp	Val	Met	Lys	Lys	Tyr	Glu	Glu	Phe	Val	Cys	Lys	Trp	Lys	Gln	Val
			580					585					590		
Gly	Gln	Pro	Lys	Leu	Phe	Phe	Ala	Thr	Met	Asp	Ile	Glu	Lys	Cys	Tyr
		595					600					605			
Asp	Ser	Val	Asn	Arg	Glu	Lys	Leu	Ser	Thr	Phe	Leu	Lys	Thr	Thr	Lys
	610					615					620				
Leu	Leu	Ser	Ser	Asp	Phe	Trp	Ile	Met	Thr	Ala	Gln	Ile	Leu	Lys	Arg
625					630					635					640
Lys	Asn	Asn	Ile	Val	Ile	Asp	Ser	Lys	Asn	Phe	Arg	Lys	Lys	Glu	Met
				645					650					655	
Lys	Asp	Tyr	Phe	Arg	Gln	Lys	Phe	Gln	Lys	Ile	Ala	Leu	Glu	Gly	Gly
			660					665					670		
Gln	Tyr	Pro	Thr	Leu	Phe	Ser	Val	Leu	Glu	Asn	Glu	Gln	Asn	Asp	Leu
		675					680					685			
Asn	Ala	Lys	Lys	Thr	Leu	Ile	Val	Glu	Ala	Lys	Gln	Arg	Asn	Tyr	Phe
	690					695					700				
Lys	Lys	Asp	Asn	Leu	Leu	Gln	Pro	Val	Ile	Asn	Ile	Cys	Gln	Tyr	Asn
705					710					715					720

Tyr Ile Asn Phe Asn Gly Lys Phe Tyr Lys Gln Thr Lys Gly Ile Pro  
 725 730 735  
 Gln Gly Leu Cys Val Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala Thr  
 740 745 750  
 Leu Glu Glu Ser Ser Leu Gly Phe Leu Arg Asp Glu Ser Met Asn Pro  
 755 760 765  
 Glu Asn Pro Asn Val Asn Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu  
 770 775 780  
 Leu Ile Thr Thr Gln Glu Asn Asn Ala Val Leu Phe Ile Glu Lys Leu  
 785 790 795 800  
 Ile Asn Val Ser Arg Glu Asn Gly Phe Lys Phe Asn Met Lys Lys Leu  
 805 810 815  
 Gln Thr Ser Phe Pro Leu Ser Pro Ser Lys Phe Ala Lys Tyr Gly Met  
 820 825 830  
 Asp Ser Val Glu Glu Gln Asn Ile Val Gln Asp Tyr Cys Asp Trp Ile  
 835 840 845  
 Gly Ile Ser Ile Asp Met Lys Thr Leu Ala Leu Met Pro Asn Ile Asn  
 850 855 860  
 Leu Arg Ile Glu Gly Ile Leu Cys Thr Leu Asn Leu Asn Met Gln Thr  
 865 870 875 880  
 Lys Lys Ala Ser Met Trp Leu Lys Lys Lys Leu Lys Ser Phe Leu Met  
 885 890 895  
 Asn Asn Ile Thr His Tyr Phe Arg Lys Thr Ile Thr Thr Glu Asp Phe  
 900 905 910  
 Ala Asn Lys Thr Leu Asn Lys Leu Phe Ile Ser Gly Gly Tyr Lys Tyr  
 915 920 925  
 Met Gln Cys Ala Lys Glu Tyr Lys Asp His Phe Lys Lys Asn Leu Ala  
 930 935 940  
 Met Ser Ser Met Ile Asp Leu Glu Val Ser Lys Ile Ile Tyr Ser Val  
 945 950 955 960  
 Thr Arg Ala Phe Phe Lys Tyr Leu Val Cys Asn Ile Lys Asp Thr Ile  
 965 970 975  
 Phe Gly Glu Glu His Tyr Pro Asp Phe Phe Leu Ser Thr Leu Lys His  
 980 985 990  
 Phe Ile Glu Ile Phe Ser Thr Lys Lys Tyr Ile Phe Asn Arg Val Cys  
 995 1000 1005  
 Met Ile Leu Lys Ala Lys Glu Ala Lys Leu Lys Ser Asp Gln Cys Gln  
 1010 1015 1020  
 Ser Leu Ile Gln Tyr Asp Ala  
 1025 1030

(2) INFORMATION FOR SEQ ID NO:111:



- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5544 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: join(959..1216, 1273..1353, 1425..1543, 1595..1857, 1894..2286, 2326..2396, 2436..2705, 2746..2862, 2914..3083, 3125..3309, 3356..3504, 3546..3759, 3797..4046, 4086..4252, 4296..4392, 4435..4597)  
 (D) OTHER INFORMATION: /note= "Schizosaccharomyces pombe telomerase catalytic subunit (TRT)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GGTACCGATT TACTTTCCTT TCTTCATAAG CTAATTGCTT CCTCGAACGC TCCTAAATCT	60
CTGGAAATAT TTTTACAAGA ACTCAATAAC AATACCAAGT CAAATTCCAA TATGAAGGTG	120
TTATTAGTGA TCGATAATAT TTCTATTTTA TCGGTCGTTA CCAAGTATAA GGACAAAAAG	180
AACAACTTCC TTCCCCCTAA AGACTTTTAC TTTATTAATT TACTTTTCAA ATATATTTTCG	240
GGTTCGCTTA CTTTAAATCG TGGTACTGTT TTAGCTGCTA CTTCTAGCCA ACCGCGTGTT	300
TCTACCCCGT CATTTGGATAT AGCTCTTGA GTAGCTCACA GAAATCCTTA CAAATCTTCT	360
GATGAGACTA TATTAGATTC ATTACAGTCC GTGCATATTC TTAACATGGA GCCTTACACT	420
TTAGATGAGT CACGTCGCAT GATGGAGTAT TTGGTATCAT CCAACGTTTG CCTTGAAAAG	480
GTTGATAATT ATTTGCAAAA TCATGTCCTT AGTGGTGGTA ATCCGCGAAA GTTTTTTGAT	540
GCTTGACACAC GTCTAGCATG ATTGAGATAT TCAAAAATTT CTATCCACTA CAACTCCTTT	600
AACGCGGTTT TATTTTTCTA TTTTCTATT TCATGTTGTT CCAAATATGT ATCATCTCGT	660
ATTAGGCTTT TTTCCGTTTT ACTCCTGGAA TCGTACCTTT TTCACTATTC CCCCTAATGA	720
ATAATCTAAA TTAGTTTCGC TTATAATTGA TAGTAGTAGA AAGATTGGTG ATTCTACTCG	780
TGTAATGTTA TTAGTTTAAA GATACTTTGC AAAACATTTA TTAGCTATCA TTATATAAAA	840
AAAATCCTAT AATTATAAAT ATTAATCAAT ATTTGCGGTC ACTATTTATT TAAAACGTTA	900
TGATCAGTAG GACACTTTGC ATATATATAG TTATGCTTAA TGGTTACTTG TAACTTGC	958
ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG	1006
Met Thr Glu His His Thr Pro Lys Ser Arg Ile Leu Arg Phe Leu Glu	
1 5 10 15	
AAT CAA TAT GTA TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT	1054
Asn Gln Tyr Val Tyr Leu Cys Thr Leu Asn Asp Tyr Val Gln Leu Val	
20 25 30	

TTG AGA GGG TCG CCG GCA AGC TCG TAT AGC AAT ATA TGC GAA CGC TTG Leu Arg Gly Ser Pro Ala Ser Ser Tyr Ser Asn Ile Cys Glu Arg Leu 35 40 45	1102
AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT CAT TCG ACT GTA Arg Ser Asp Val Gln Thr Ser Phe Ser Ile Phe Leu His Ser Thr Val 50 55 60	1150
GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA Val Gly Phe Asp Ser Lys Pro Asp Glu Gly Val Gln Phe Ser Ser Pro 65 70 75 80	1198
AAA TGC TCA CAG TCA GAG GTATATATAT TTTTGTTTTG ATTTTTTTTCT Lys Cys Ser Gln Ser Glu 85	1246
ATTCGGGATA GCTAATATAT GGGCAG CTA ATA GCG AAT GTT GTA AAA CAG ATG Leu Ile Ala Asn Val Val Lys Gln Met 90 95	1299
TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA CTG ATG AAA GGG TTT Phe Asp Glu Ser Phe Glu Arg Arg Arg Asn Leu Leu Met Lys Gly Phe 100 105 110	1347
TCC ATG GTAAGGTATT CTAATTGTGA AATATTTACC TGCAATTACT GTTTCAAAGA Ser Met	1403
GATTGTATTT AACCGATAAA G AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA Asn His Glu Asp Phe Arg Ala Met His Val 115 120	1454
AAC GGA GTA CAA AAT GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA Asn Gly Val Gln Asn Asp Leu Val Ser Thr Phe Pro Asn Tyr Leu Ile 125 130 135	1502
TCT ATA CTT GAG TCA AAA AAT TGG CAA CTT TTG TTA GAA AT Ser Ile Leu Glu Ser Lys Asn Trp Gln Leu Leu Glu Ile 140 145 150	1543
GTAAATACCG GTTAAGATGT TCGCACTTT GAACAAGACT GACAAGTATA G T ATC Ile	1598
GGC AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG Gly Ser Asp Ala Met His Tyr Leu Leu Ser Lys Gly Ser Ile Phe Glu 155 160 165 170	1646
GCT CTT CCA AAT GAC AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT Ala Leu Pro Asn Asp Asn Tyr Leu Gln Ile Ser Gly Ile Pro Leu Phe 175 180 185	1694
AAA AAT AAT GTG TTT GAG GAA ACT GTG TCA AAA AAA AGA AAG CGA ACC Lys Asn Asn Val Phe Glu Glu Thr Val Ser Lys Lys Arg Lys Arg Thr 190 195 200	1742
ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA GAA GTT TCC Ile Glu Thr Ser Ile Thr Gln Asn Lys Ser Ala Arg Lys Glu Val Ser 205 210 215	1790
TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC Trp Asn Ser Ile Ser Ile Ser Arg Phe Ser Ile Phe Tyr Arg Ser Ser	1838

220	225	230	
TAT AAG AAG TTT AAG CAA G	GTA	CTA	1887
Tyr Lys Lys Phe Lys Gln	CTG	TAT	
235	240	CTC TAT	
TTTTAG	AT	CTA	1934
Asp Leu Tyr Phe Asn Leu His Ser Ile Cys Asp Arg Asn Thr			
	245	250	
GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG CAA TTT GGA CTT ATA			1982
Val His Met Trp Leu Gln Trp Ile Phe Pro Arg Gln Phe Gly Leu Ile			
255	260	265	270
AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA TCA			2030
Asn Ala Phe Gln Val Lys Gln Leu His Lys Val Ile Pro Leu Val Ser			
	275	280	285
CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT			2078
Gln Ser Thr Val Val Pro Lys Arg Leu Leu Lys Val Tyr Pro Leu Ile			
	290	295	300
GAA CAA ACA GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC			2126
Glu Gln Thr Ala Lys Arg Leu His Arg Ile Ser Leu Ser Lys Val Tyr			
	305	310	315
AAC CAT TAT TGC CCA TAT ATT GAC ACC CAC GAT GAT GAA AAA ATC CTT			2174
Asn His Tyr Cys Pro Tyr Ile Asp Thr His Asp Asp Glu Lys Ile Leu			
	320	325	330
AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG TTT CTT CGA TCC ATT			2222
Ser Tyr Ser Leu Lys Pro Asn Gln Val Phe Ala Phe Leu Arg Ser Ile			
	335	340	350
CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA TTT			2270
Leu Val Arg Val Phe Pro Lys Leu Ile Trp Gly Asn Gln Arg Ile Phe			
	355	360	365
GAG ATA ATA TTA AAA G	GTATTGTATA	AAATTTATTA	2327
Glu Ile Ile Leu Lys		CCACTAACGA	Asp
	370	TTTTACCAG	
CTC GAA ACT TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT			2375
Leu Glu Thr Phe Leu Lys Leu Ser Arg Tyr Glu Ser Phe Ser Leu His			
	375	380	385
TAT TTA ATG AGT AAC ATA AAG GTAATATGCC AAATTTTTTTT ACCATTAATT			2426
Tyr Leu Met Ser Asn Ile Lys			
	390	395	
AACAATCAG ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA AAA AGG TCA			2474
Ile Ser Glu Ile Glu Trp Leu Val Leu Gly Lys Arg Ser			
	400	405	
AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT			2522
Asn Ala Lys Met Cys Leu Ser Asp Phe Glu Lys Arg Lys Gln Ile Phe			
	410	415	420
GCG GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA			2570
Ala Glu Phe Ile Tyr Trp Leu Tyr Asn Ser Phe Ile Ile Pro Ile Leu			
	425	430	435
			440

CAA TCT TTT TTT TAT ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT	2618
Gln Ser Phe Phe Tyr Ile Thr Glu Ser Ser Asp Leu Arg Asn Arg Thr	
445 450 455	
GTT TAT TTT AGA AAA GAT ATT TGG AAA CTC TTG TGC CGA CCC TTT ATT	2666
Val Tyr Phe Arg Lys Asp Ile Trp Lys Leu Leu Cys Arg Pro Phe Ile	
460 465 470	
ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG GTATTTTAAA	2715
Thr Ser Met Lys Met Glu Ala Phe Glu Lys Ile Asn Glu	
475 480 485	
GTATTTTTTG CAAAAAGCTA ATATTTTTCAG AAC AAT GTT AGG ATG GAT ACT CAG	2769
Asn Asn Val Arg Met Asp Thr Gln	
490	
AAA ACT ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT	2817
Lys Thr Thr Leu Pro Pro Ala Val Ile Arg Leu Leu Pro Lys Lys Asn	
495 500 505	
ACC TTT CGT CTC ATT ACG AAT TTA AGA AAA AGA TTC TTA ATA AAG	2862
Thr Phe Arg Leu Ile Thr Asn Leu Arg Lys Arg Phe Leu Ile Lys	
510 515 520	
GTATTAATTT TTGGTCATCA ATGTACTTTA CTTCTAATCT ATTATTAGCA G ATG GGT	2919
Met Gly	
525	
TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG	2967
Ser Asn Lys Lys Met Leu Val Ser Thr Asn Gln Thr Leu Arg Pro Val	
530 535 540	
GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA	3015
Ala Ser Ile Leu Lys His Leu Ile Asn Glu Glu Ser Ser Gly Ile Pro	
545 550 555	
TTT AAC TTG GAG GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT	3063
Phe Asn Leu Glu Val Tyr Met Lys Leu Leu Thr Phe Lys Lys Asp Leu	
560 565 570	
CTT AAG CAC CGA ATG TTT GG GTAATTATAT AATGCGCGAT TCCTCATTAT	3113
Leu Lys His Arg Met Phe Gly	
575 580	
TAATTTTGCA G G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA AAA TCC	3161
Arg Lys Lys Tyr Phe Val Arg Ile Asp Ile Lys Ser	
585 590	
TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG	3209
Cys Tyr Asp Arg Ile Lys Gln Asp Leu Met Phe Arg Ile Val Lys Lys	
595 600 605	
AAA CTC AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT	3257
Lys Leu Lys Asp Pro Glu Phe Val Ile Arg Lys Tyr Ala Thr Ile His	
610 615 620 625	
GCA ACA AGT GAC CGA GCT ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC	3305
Ala Thr Ser Asp Arg Ala Thr Lys Asn Phe Val Ser Glu Ala Phe Ser	
630 635 640	
TAT T GTAAGTTTAT TTTTTCATTG GAATTTTTTA ACAAATTCTT TTTTAG TT	3357
Tyr Phe	

GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA Asp Met Val Pro Phe Glu Lys Val Val Gln Leu Leu Ser Met Lys Thr 645 650 655	3405
TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT Ser Asp Thr Leu Phe Val Asp Phe Val Asp Tyr Trp Thr Lys Ser Ser 660 665 670 675	3453
TCT GAA ATT TTT AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT Ser Glu Ile Phe Lys Met Leu Lys Glu His Leu Ser Gly His Ile Val 680 685 690	3501
AAG GTATACCAAT TGTGGAATTG TAATAACACT AATGAACTA G ATA GGA AAT Lys Ile Gly Asn 695	3554
TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA ATT CTG TCA Ser Gln Tyr Leu Gln Lys Val Gly Ile Pro Gln Gly Ser Ile Leu Ser 700 705 710	3602
TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA Ser Phe Leu Cys His Phe Tyr Met Glu Asp Leu Ile Asp Glu Tyr Leu 715 720 725	3650
TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT Ser Phe Thr Lys Lys Lys Gly Ser Val Leu Leu Arg Val Val Asp Asp 730 735 740	3698
TTC CTC TTT ATA ACA GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT Phe Leu Phe Ile Thr Val Asn Lys Lys Asp Ala Lys Lys Phe Leu Asn 745 750 755	3746
TTA TCT TTA AGA G GTGAGTTGCT GTCATTCCTA AGTTCTAACC GTTGAAG GA Leu Ser Leu Arg Gly 760	3798
TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA ATA AAC Phe Glu Lys His Asn Phe Ser Thr Ser Leu Glu Lys Thr Val Ile Asn 765 770 775 780	3846
TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC Phe Glu Asn Ser Asn Gly Ile Ile Asn Asn Thr Phe Phe Asn Glu Ser 785 790 795	3894
AAG AAA AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT Lys Lys Arg Met Pro Phe Phe Gly Phe Ser Val Asn Met Arg Ser Leu 800 805 810	3942
GAT ACA TTG TTA GCA TGT CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT Asp Thr Leu Leu Ala Cys Pro Lys Ile Asp Glu Ala Leu Phe Asn Ser 815 820 825	3990
ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG AAA TCT TTT TTT TAC AAA Thr Ser Val Glu Leu Thr Lys His Met Gly Lys Ser Phe Phe Tyr Lys 830 835 840	4038
ATT CTA AG GTATACTGTG TAACTGAATA ATAGCTGACA AATAATCAG A TCG Ile Leu Arg Ser 845	4089

AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA Ser Leu Ala Ser Phe Ala Gln Val Phe Ile Asp Ile Thr His Asn Ser 850 855 860	4137
AAA TTC AAT TCT TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT Lys Phe Asn Ser Cys Cys Asn Ile Tyr Arg Leu Gly Tyr Ser Met Cys 865 870 875 880	4185
ATG AGA GCA CAA GCA TAC TTA AAA AGG ATG AAG GAT ATA TTT ATT CCC Met Arg Ala Gln Ala Tyr Leu Lys Arg Met Lys Asp Ile Phe Ile Pro 885 890 895	4233
CAA AGA ATG TTC ATA ACG G GTGAGTACTT ATTTTAACTA GAAAAGTCAT Gln Arg Met Phe Ile Thr 900	4282
TAATTAACCT TAG AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA Asp Leu Leu Asn Val Ile Gly Arg Lys Ile Trp Lys 905 910	4330
AAG TTG GCC GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT Lys Leu Ala Glu Ile Leu Gly Tyr Thr Ser Arg Arg Phe Leu Ser Ser 915 920 925 930	4378
GCA GAA GTC AAA TG GTACGTGTG GTCTCGAGAC TTCAGCAATA TTGACACATC Ala Glu Val Lys Trp 935	4432
AG G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA CCC TCT TTC AAA Leu Phe Cys Leu Gly Met Arg Asp Gly Leu Lys Pro Ser Phe Lys 940 945 950	4480
TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT Tyr His Pro Cys Phe Glu Gln Leu Ile Tyr Gln Phe Gln Ser Leu Thr 955 960 965	4528
GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA Asp Leu Ile Lys Pro Leu Arg Pro Val Leu Arg Gln Val Leu Phe Leu 970 975 980	4576
CAT AGA AGA ATA GCT GAT TAATGTCATT TTCAATTTAT TATATACATC His Arg Arg Ile Ala Asp 985	4624
CTTTATTACT GGTGTCTTAA ACAATATTAT TACTAAGTAT AGCTGACCCC CAAAGCAAGC	4684
ATACTATAGG ATTCTAGTA AAGTAAAATT AATCTCGTTA TTAGTTTTGA TTGACTTGTC	4744
TTTATCCTTA TACTTTTAAAG AAAGATTGAC AGTGGTTGCT GACTACTGCC CACATGCCCA	4804
TTAAACGGGA GTGGTTAAAC ATTAAAAGTA ATACATGAGG CTAATCTCCT TTCATTTAGA	4864
ATAAGGAAAG TGGTTTTCTA TAATGAATAA TGCCCGCACT AATGCAAAAA GACGAAGATT	4924
ATCTTCTAAA CAAGGGGGAT TAAGCATATC CGAAGGAAAA GAGAGTAATA TACCCAGTGT	4984
TGTTGAAGAA AGCAAGGATA ATTTGGAACA AGCTTCTGCA GATGACAGGC TAAATTTTGG	5044
TGACCGAATT TTGGTAAAAG CCCCAGGTTA TCCATGGTGG CCGGCCTTGC TACTGAGACG	5104
AAAAGAACT AAGGATAGTT TGAATACTAA TAGCTCATTT AATGTCTTAT ATAAGGTTTT	5164

GTTTTTTCCT GACTTCAATT TTGCATGGGT GAAAAGAAAT AGTGTTAAGC CATTATTGGA	5224
TTCCGAAATA GCCAAATTC TTGGTTCCTC AAAGCGGAAG TCTAAAGAAC TTATTGAAGC	5284
TTATGAGGCT TCAAAACTC CTCCTGATTT AAAGGAGGAA TCTTCCACCG ATGAGGAAAT	5344
GGATAGCTTA TCAGCTGCTG AGGAGAAGCC TAATTTTTTG CAAAAAGAA AATATCATTG	5404
GGAGACATCT CTTGATGAAT CAGATGCGGA GAGTATCTCC AGCGGATCCT TGATGTCAAT	5464
AACTTCTATT TCTGAAATGT ATGGTCCTAC TGTCGCTTCG ACTTCTCGTA GCTCTACGCA	5524
GTAAAGTGAC CAAAGGTACC	5544

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 988 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Met	Thr	Glu	His	His	Thr	Pro	Lys	Ser	Arg	Ile	Leu	Arg	Phe	Leu	Glu	1	5	10	15
Asn	Gln	Tyr	Val	Tyr	Leu	Cys	Thr	Leu	Asn	Asp	Tyr	Val	Gln	Leu	Val	20	25	30	
Leu	Arg	Gly	Ser	Pro	Ala	Ser	Ser	Tyr	Ser	Asn	Ile	Cys	Glu	Arg	Leu	35	40	45	
Arg	Ser	Asp	Val	Gln	Thr	Ser	Phe	Ser	Ile	Phe	Leu	His	Ser	Thr	Val	50	55	60	
Val	Gly	Phe	Asp	Ser	Lys	Pro	Asp	Glu	Gly	Val	Gln	Phe	Ser	Ser	Pro	65	70	75	80
Lys	Cys	Ser	Gln	Ser	Glu	Leu	Ile	Ala	Asn	Val	Val	Lys	Gln	Met	Phe	85	90	95	
Asp	Glu	Ser	Phe	Glu	Arg	Arg	Arg	Asn	Leu	Leu	Met	Lys	Gly	Phe	Ser	100	105	110	
Met	Asn	His	Glu	Asp	Phe	Arg	Ala	Met	His	Val	Asn	Gly	Val	Gln	Asn	115	120	125	
Asp	Leu	Val	Ser	Thr	Phe	Pro	Asn	Tyr	Leu	Ile	Ser	Ile	Leu	Glu	Ser	130	135	140	
Lys	Asn	Trp	Gln	Leu	Leu	Leu	Glu	Ile	Ile	Gly	Ser	Asp	Ala	Met	His	145	150	155	160
Tyr	Leu	Leu	Ser	Lys	Gly	Ser	Ile	Phe	Glu	Ala	Leu	Pro	Asn	Asp	Asn	165	170	175	
Tyr	Leu	Gln	Ile	Ser	Gly	Ile	Pro	Leu	Phe	Lys	Asn	Asn	Val	Phe	Glu	180	185	190	

Glu Thr Val Ser Lys Lys Arg Lys Arg Thr Ile Glu Thr Ser Ile Thr  
 195 200 205  
 Gln Asn Lys Ser Ala Arg Lys Glu Val Ser Trp Asn Ser Ile Ser Ile  
 210 215 220  
 Ser Arg Phe Ser Ile Phe Tyr Arg Ser Ser Tyr Lys Lys Phe Lys Gln  
 225 230 235 240  
 Asp Leu Tyr Phe Asn Leu His Ser Ile Cys Asp Arg Asn Thr Val His  
 245 250 255  
 Met Trp Leu Gln Trp Ile Phe Pro Arg Gln Phe Gly Leu Ile Asn Ala  
 260 265 270  
 Phe Gln Val Lys Gln Leu His Lys Val Ile Pro Leu Val Ser Gln Ser  
 275 280 285  
 Thr Val Val Pro Lys Arg Leu Leu Lys Val Tyr Pro Leu Ile Glu Gln  
 290 295 300  
 Thr Ala Lys Arg Leu His Arg Ile Ser Leu Ser Lys Val Tyr Asn His  
 305 310 315 320  
 Tyr Cys Pro Tyr Ile Asp Thr His Asp Asp Glu Lys Ile Leu Ser Tyr  
 325 330 335  
 Ser Leu Lys Pro Asn Gln Val Phe Ala Phe Leu Arg Ser Ile Leu Val  
 340 345 350  
 Arg Val Phe Pro Lys Leu Ile Trp Gly Asn Gln Arg Ile Phe Glu Ile  
 355 360 365  
 Ile Leu Lys Asp Leu Glu Thr Phe Leu Lys Leu Ser Arg Tyr Glu Ser  
 370 375 380  
 Phe Ser Leu His Tyr Leu Met Ser Asn Ile Lys Ile Ser Glu Ile Glu  
 385 390 395 400  
 Trp Leu Val Leu Gly Lys Arg Ser Asn Ala Lys Met Cys Leu Ser Asp  
 405 410 415  
 Phe Glu Lys Arg Lys Gln Ile Phe Ala Glu Phe Ile Tyr Trp Leu Tyr  
 420 425 430  
 Asn Ser Phe Ile Ile Pro Ile Leu Gln Ser Phe Phe Tyr Ile Thr Glu  
 435 440 445  
 Ser Ser Asp Leu Arg Asn Arg Thr Val Tyr Phe Arg Lys Asp Ile Trp  
 450 455 460  
 Lys Leu Leu Cys Arg Pro Phe Ile Thr Ser Met Lys Met Glu Ala Phe  
 465 470 475 480  
 Glu Lys Ile Asn Glu Asn Asn Val Arg Met Asp Thr Gln Lys Thr Thr  
 485 490 495  
 Leu Pro Pro Ala Val Ile Arg Leu Leu Pro Lys Lys Asn Thr Phe Arg  
 500 505 510  
 Leu Ile Thr Asn Leu Arg Lys Arg Phe Leu Ile Lys Met Gly Ser Asn  
 515 520 525



Lys Lys Met Leu Val Ser Thr Asn Gln Thr Leu Arg Pro Val Ala Ser  
 530 535 540  
 Ile Leu Lys His Leu Ile Asn Glu Glu Ser Ser Gly Ile Pro Phe Asn  
 545 550 555 560  
 Leu Glu Val Tyr Met Lys Leu Leu Thr Phe Lys Lys Asp Leu Leu Lys  
 565 570 575  
 His Arg Met Phe Gly Arg Lys Lys Tyr Phe Val Arg Ile Asp Ile Lys  
 580 585 590  
 Ser Cys Tyr Asp Arg Ile Lys Gln Asp Leu Met Phe Arg Ile Val Lys  
 595 600 605  
 Lys Lys Leu Lys Asp Pro Glu Phe Val Ile Arg Lys Tyr Ala Thr Ile  
 610 615 620  
 His Ala Thr Ser Asp Arg Ala Thr Lys Asn Phe Val Ser Glu Ala Phe  
 625 630 635 640  
 Ser Tyr Phe Asp Met Val Pro Phe Glu Lys Val Val Gln Leu Leu Ser  
 645 650 655  
 Met Lys Thr Ser Asp Thr Leu Phe Val Asp Phe Val Asp Tyr Trp Thr  
 660 665 670  
 Lys Ser Ser Ser Glu Ile Phe Lys Met Leu Lys Glu His Leu Ser Gly  
 675 680 685  
 His Ile Val Lys Ile Gly Asn Ser Gln Tyr Leu Gln Lys Val Gly Ile  
 690 695 700  
 Pro Gln Gly Ser Ile Leu Ser Ser Phe Leu Cys His Phe Tyr Met Glu  
 705 710 715 720  
 Asp Leu Ile Asp Glu Tyr Leu Ser Phe Thr Lys Lys Lys Gly Ser Val  
 725 730 735  
 Leu Leu Arg Val Val Asp Asp Phe Leu Phe Ile Thr Val Asn Lys Lys  
 740 745 750  
 Asp Ala Lys Lys Phe Leu Asn Leu Ser Leu Arg Gly Phe Glu Lys His  
 755 760 765  
 Asn Phe Ser Thr Ser Leu Glu Lys Thr Val Ile Asn Phe Glu Asn Ser  
 770 775 780  
 Asn Gly Ile Ile Asn Asn Thr Phe Phe Asn Glu Ser Lys Lys Arg Met  
 785 790 795 800  
 Pro Phe Phe Gly Phe Ser Val Asn Met Arg Ser Leu Asp Thr Leu Leu  
 805 810 815  
 Ala Cys Pro Lys Ile Asp Glu Ala Leu Phe Asn Ser Thr Ser Val Glu  
 820 825 830  
 Leu Thr Lys His Met Gly Lys Ser Phe Phe Tyr Lys Ile Leu Arg Ser  
 835 840 845  
 Ser Leu Ala Ser Phe Ala Gln Val Phe Ile Asp Ile Thr His Asn Ser  
 850 855 860

Lys Phe Asn Ser Cys Cys Asn Ile Tyr Arg Leu Gly Tyr Ser Met Cys  
 865 870 875 880  
 Met Arg Ala Gln Ala Tyr Leu Lys Arg Met Lys Asp Ile Phe Ile Pro  
 885 890 895  
 Gln Arg Met Phe Ile Thr Asp Leu Leu Asn Val Ile Gly Arg Lys Ile  
 900 905 910  
 Trp Lys Lys Leu Ala Glu Ile Leu Gly Tyr Thr Ser Arg Arg Phe Leu  
 915 920 925  
 Ser Ser Ala Glu Val Lys Trp Leu Phe Cys Leu Gly Met Arg Asp Gly  
 930 935 940  
 Leu Lys Pro Ser Phe Lys Tyr His Pro Cys Phe Glu Gln Leu Ile Tyr  
 945 950 955 960  
 Gln Phe Gln Ser Leu Thr Asp Leu Ile Lys Pro Leu Arg Pro Val Leu  
 965 970 975  
 Arg Gln Val Leu Phe Leu His Arg Arg Ile Ala Asp  
 980 985

(2) INFORMATION FOR SEQ ID NO:113:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Phe Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe  
 1 5 10 15  
 Tyr Arg Lys Ser Val Trp Ser Lys  
 20

(2) INFORMATION FOR SEQ ID NO:114:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Arg Gln His Leu Lys Arg Val Gln Leu Arg Asp Val Ser Glu Ala Glu  
 1 5 10 15  
 Val Arg Gln His Arg Glu Ala

## (2) INFORMATION FOR SEQ ID NO:115:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

Ala	Arg	Thr	Phe	Arg	Arg	Glu	Lys	Arg	Ala	Glu	Arg	Leu	Thr	Ser	Arg
1				5					10					15	
Val	Lys	Ala	Leu	Phe	Ser	Val	Leu	Asn	Tyr	Glu					
			20					25							

## (2) INFORMATION FOR SEQ ID NO:116:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

Pro	Ala	Leu	Leu	Thr	Ser	Arg	Leu	Arg	Phe	Ile	Pro	Lys	Pro	Asp	Gly
1				5					10					15	
Leu	Arg	Pro	Ile	Val	Asn	Met	Asp	Tyr	Val	Val					
			20					25							

## (2) INFORMATION FOR SEQ ID NO:117:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 2
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Leu or Ile"

## (ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 7..8

(ix) FEATURE:

(B) LOCATION: 10..11

```
/product="GlnXaa"  
/note="Xaa = hydrophobic amino acid,  
Ala, Leu, Ile, Val, Pro, Phe, Trp  
or Met"
```

(A) NAME/KEY: Modified-site

(B) LOCATION: 12

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Gln or Arg"

FEATURE:  
(A) NAME/KEY: Modified-site

(B) LOCATION: 13

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = polar amino acid, Gly,  
Ser, Thr, Tyr, Cys, Asn or Gln"

FEATURE:  
(A) NAME/KEY: Modified-site

(B) LOCATION: 21

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = polar amino acid, Gly,  
Ser, Thr, Tyr, Cys, Asn or Gln"

(A) NAME/KEY: Modified-site

(B) LOCATION: 25

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = polar amino acid, Gly,  
Ser, Thr, Tyr, Cys, Asn or Gln"

FEATURE:  
(A) NAME/KEY: Modified-site

(B) LOCATION: 28..29

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Phe or Tyr"

(A) NAME/KEY: Modified-site

(B) LOCATION: 31

(D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Lys or His"

Trp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Phe Tyr  
1 5 10 15

Xaa Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Xaa Xaa  
20 25 30

91

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 2
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Leu or Ile"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 7..8
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid,  
Ala, Leu, Ile, Val, Pro, Phe, Trp  
or Met"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 10..11
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = hydrophobic amino acid,  
Ala, Leu, Ile, Val, Pro, Phe, Trp  
or Met"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 12
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Gln or Arg"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 13
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = polar amino acid, Gly,  
Ser, Thr, Tyr, Cys, Asn or Gln"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 21
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = polar amino acid, Gly,  
Ser, Thr, Tyr, Cys, Asn or Gln"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 25
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = polar amino acid, Gly,  
Ser, Thr, Tyr, Cys, Asn or Gln"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 29..30
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Phe or Tyr"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 32
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Lys or His"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

Trp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Phe Tyr  
1 5 10 15  
Xaa Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Xaa  
20 25 30  
Xaa Xaa Trp  
35

(2) INFORMATION FOR SEQ ID NO:119:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Trp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Phe Tyr  
1 5 10 15  
Xaa Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Xaa Xaa  
20 25 30  
Xaa Trp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ile  
35 40

(2) INFORMATION FOR SEQ ID NO:120:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

Trp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Phe Tyr

1	5	10	15
Xaa Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Xaa			
20	25	30	
Xaa Xaa Trp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ile			
35	40		

(2) INFORMATION FOR SEQ ID NO:121:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

Glu Xaa Xaa Val Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:122:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 11 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Xaa Xaa Xaa Arg Xaa Xaa Pro Lys Xaa Xaa Xaa
1 5 10

(2) INFORMATION FOR SEQ ID NO:123:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Xaa Arg Xaa Ile Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:124:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 17 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Tyr Asp Xaa  
1 5 10 15  
Xaa

(2) INFORMATION FOR SEQ ID NO:125:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Tyr Xaa Xaa Xaa Xaa Gly Xaa Xaa Gln Gly Xaa Xaa Xaa Ser Xaa Xaa  
1 5 10 15  
Xaa Xaa Xaa Xaa Xaa Xaa  
20

(2) INFORMATION FOR SEQ ID NO:126:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 13 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

Xaa Xaa Xaa Xaa Xaa Xaa Asp Asp Xaa Leu Xaa Xaa Xaa  
1 5 10

(2) INFORMATION FOR SEQ ID NO:127:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear



(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

Phe Phe Tyr Xaa Thr Glu  
1 5

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Phe Phe Tyr Val Thr Glu  
1 5

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

TTYTTYTAYG TNACNGA

17

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

TCNGTNACRT ARAARAA

17

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Arg Phe Ile Pro Lys Pro  
1 5

(2) INFORMATION FOR SEQ ID NO:132:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

MGNTTYATHC CNAARCC

17

(2) INFORMATION FOR SEQ ID NO:133:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

GGYTTNGGDA TRAANC

16

(2) INFORMATION FOR SEQ ID NO:134:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

Ala Tyr Asp Thr Ile  
1 5

(2) INFORMATION FOR SEQ ID NO:135:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

GCNTAYGAYA CNAT

14

(2) INFORMATION FOR SEQ ID NO:136:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

TANGTRTCRT ANGC

14

(2) INFORMATION FOR SEQ ID NO:137:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

Gly Ile Pro Gln Gly  
1 5

(2) INFORMATION FOR SEQ ID NO:138:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

## (2) INFORMATION FOR SEQ ID NO:139:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

SWNCCYTGNNG GDATNCC

17

## (2) INFORMATION FOR SEQ ID NO:140:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

Leu	Val	Asp	Asp	Phe	Leu
1				5	

## (2) INFORMATION FOR SEQ ID NO:141:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

YTNGTNGAYG AYTYYT

17

## (2) INFORMATION FOR SEQ ID NO:142:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

Asp Asp Phe Leu Leu Val Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:143:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

GTNACNARNA RRAARTCRTC

20

(2) INFORMATION FOR SEQ ID NO:144:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

GTGAAGGCAC TGTTTCAGCG

19

(2) INFORMATION FOR SEQ ID NO:145:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

CGCGTGGGTG AGGTGAGGTG

20

(2) INFORMATION FOR SEQ ID NO:146:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

CTGTGCTGGG CCTGGACGAT A

21

(2) INFORMATION FOR SEQ ID NO:147:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

AGCTTGTTCT CCATGTCGCC GTAG

24

(2) INFORMATION FOR SEQ ID NO:148:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

GTGGATGATT TCTTGTTGG

19

(2) INFORMATION FOR SEQ ID NO:149:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CTGGACACTC AGCCCTTGG

19

(2) INFORMATION FOR SEQ ID NO:150:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

GGCAGGTGTG CTGGACACT

19

(2) INFORMATION FOR SEQ ID NO:151:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

TTTGATGATG CTGGCGATG

19

(2) INFORMATION FOR SEQ ID NO:152:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

GGGGCTCGTC TTCTACAGG

19

(2) INFORMATION FOR SEQ ID NO:153:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

CAGCAGGAGG ATCTTGATG

19

(2) INFORMATION FOR SEQ ID NO:154:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

TGACCCCAGG AGTGGCACG

19

(2) INFORMATION FOR SEQ ID NO:155:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

TCAAGCTGAC TCGACACCG

19

(2) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

CGGCGTGACA GGGCTGC

17

(2) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

GCTGAAGGCT GAGTGTCC

18

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs



- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

TAGTCCATGT TCACAATCG

19

(2) INFORMATION FOR SEQ ID NO:159:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

TTTCCGTGTT GAGTGTTTC

19

(2) INFORMATION FOR SEQ ID NO:160:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

GTCACCGTGT TGGGCAGG

18

(2) INFORMATION FOR SEQ ID NO:161:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

GCTACCTGCC CAACACGG

18

(2) INFORMATION FOR SEQ ID NO:162:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

GCGCGAAGAA CGTGCTGG

18

(2) INFORMATION FOR SEQ ID NO:163:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

CACTGCTCCT TGTCGCCTG

19

(2) INFORMATION FOR SEQ ID NO:164:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

TTCCCAAGGA CTTTGTTC

19

(2) INFORMATION FOR SEQ ID NO:165:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

TGTTCTCAA GACGCACTG

19

(2) INFORMATION FOR SEQ ID NO:166:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

TACTGCGTGC GTCGGTATG

19

(2) INFORMATION FOR SEQ ID NO:167:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

GGTCTTGCGG CTGAAGTGT

19

(2) INFORMATION FOR SEQ ID NO:168:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

TGGTTCACCT GCTGGCACG

19

(2) INFORMATION FOR SEQ ID NO:169:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

GTGGTTTCTG TGTGGTGTG

19

(2) INFORMATION FOR SEQ ID NO:170:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

GACACCACAC AGAAACCAC

19

(2) INFORMATION FOR SEQ ID NO:171:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

GTGCCAGCAG GTGAACCAG

19

(2) INFORMATION FOR SEQ ID NO:172:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

GCAGTGCGTC TTGAGGAGC

19

(2) INFORMATION FOR SEQ ID NO:173:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

TGGAACCATA GCGTCAGGGA G

21

(2) INFORMATION FOR SEQ ID NO:174:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

GGCCTCCCTG ACGCTATGGT T

21

(2) INFORMATION FOR SEQ ID NO:175:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

GCKCGGCGCT GCCACTCAGG

20

(2) INFORMATION FOR SEQ ID NO:176:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

GCTCGGCGCT GCCACTCAGG

20

(2) INFORMATION FOR SEQ ID NO:177:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

ACGCCGAGAC CAAGCACTTC

20

(2) INFORMATION FOR SEQ ID NO:178:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

CCAAAGAGGT GGCTTCTTCG

20

(2) INFORMATION FOR SEQ ID NO:179:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

AAGGCCAGCA CGTTCTTCGC

20

(2) INFORMATION FOR SEQ ID NO:180:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

CACGTTCTGTG CGGCGCCTG

19

(2) INFORMATION FOR SEQ ID NO:181:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

CCTTCACCAC CAGCGTGCG

19

(2) INFORMATION FOR SEQ ID NO:182:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

GGCGACGACG TGCTGGTTC

19

(2) INFORMATION FOR SEQ ID NO:183:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

GGCTCAGGGG CAGCGCCAC

19

(2) INFORMATION FOR SEQ ID NO:184:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

CTGGCAGGTG TACGGCTTC

19

(2) INFORMATION FOR SEQ ID NO:185:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

GCGTGGACCG AGTGACCGTG GTTTC

25

(2) INFORMATION FOR SEQ ID NO:186:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

GACGTGGTGG CCGCGATGTG G

21

(2) INFORMATION FOR SEQ ID NO:187:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

GAAGTCTGCC GTTGCCCAAG AG

22

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

GACACCACAC AGAAACCACG GTCAC

25

(2) INFORMATION FOR SEQ ID NO:189:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single



(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

CGCCCCCTCC TTCCGCCAGG T 21

(2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

CGAAGCCGAA GGCCAGCACG TTCTT 25

(2) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

GGTGGCCCGA GTGCTGCAGA GG 22

(2) INFORMATION FOR SEQ ID NO:192:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

GTAGCTGCGC ACGCTGGTGG TGAAG 25

(2) INFORMATION FOR SEQ ID NO:193:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

TGGGCGACGA CGTGCTGGTT CA

22

(2) INFORMATION FOR SEQ ID NO:194:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

TATGGTTCCA GGCCCGTTTCG CATCC

25

(2) INFORMATION FOR SEQ ID NO:195:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

CCAGCTGCGC CTACCAGGTG TGC

23

(2) INFORMATION FOR SEQ ID NO:196:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

GGCCTCCCTG ACGCTATGGT TCCAG

25

(2) INFORMATION FOR SEQ ID NO:197:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:  
 GGTGCTGCCG CTGGCCACGT TCG 23

(2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:  
 TCCCAGGGCA CGCACACCAG GCACT 25

(2) INFORMATION FOR SEQ ID NO:199:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:  
 GTACAGGGCA CACCTTTGGT CACTC 25

(2) INFORMATION FOR SEQ ID NO:200:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:  
 TCGACGACGT ACACACTCAT CAGCC 25

(2) INFORMATION FOR SEQ ID NO:201:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

AGCGGCAGCA CCTCGCGGTA GTGGC

25

(2) INFORMATION FOR SEQ ID NO:202:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

CCACCAGCTC CTTCAGGCAG GACAC

25

(2) INFORMATION FOR SEQ ID NO:203:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

CCAGGGCTTC CCACGTGCGC AGCAG

25

(2) INFORMATION FOR SEQ ID NO:204:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

CGCACGAACG TGGCCAGCGG CAGCA

25

(2) INFORMATION FOR SEQ ID NO:205:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

TGACCGTGGT TTCTGTGTGG TGT

23

(2) INFORMATION FOR SEQ ID NO:206:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

CCCTCTTCAA GTGCTGTCTG ATTCC

25

(2) INFORMATION FOR SEQ ID NO:207:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

ATCGCGGCCA CCACGTCCCT

20

(2) INFORMATION FOR SEQ ID NO:208:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

TGCTCCAGAC ACTCGGCCGG TAGAA

25

(2) INFORMATION FOR SEQ ID NO:209:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

ACGAAGCCGT ACACCTGCC

19

(2) INFORMATION FOR SEQ ID NO:210:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

CGACATCCCT GCGTTCTTGG CTTTC

25

(2) INFORMATION FOR SEQ ID NO:211:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

CACTGCTGGC CTCATTCAGG G

21

(2) INFORMATION FOR SEQ ID NO:212:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:  
CACTGCTGGC CTCATTCAGG G

21

(2) INFORMATION FOR SEQ ID NO:213:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:  
GCAGCCATAC TCAGGGACAC

20

(2) INFORMATION FOR SEQ ID NO:214:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:  
CCATCCTCTC CACGCTGCTC

20

(2) INFORMATION FOR SEQ ID NO:215:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:  
GCGATGACCT CCGTGAGCCT G

21

(2) INFORMATION FOR SEQ ID NO:216:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:  
CCCAGGACAG GCTCACGGA

19

(2) INFORMATION FOR SEQ ID NO:217:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:  
CCTCTTCAAG TGCTGTCTGA TTCC

24

(2) INFORMATION FOR SEQ ID NO:218:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:  
CAGCTCGACG ACGTACACAC TCATC

25

(2) INFORMATION FOR SEQ ID NO:219:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:  
CTGACGTCCA GACTCCGCTT CAT

23

(2) INFORMATION FOR SEQ ID NO:220:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single



(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:  
GACCTGAGCA GCTCGACGAC GTACACACTC ATC 33

(2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:  
GTCGTCGAGC TGCTCAGGTC 20

(2) INFORMATION FOR SEQ ID NO:222:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:  
AGCACGCTGA ACAGTGCCTT 20

(2) INFORMATION FOR SEQ ID NO:223:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:  
GACCTGAGCA GCTCGACGAC 20

(2) INFORMATION FOR SEQ ID NO:224:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

AAGGCACTGT TCAGCGTGCT

20

(2) INFORMATION FOR SEQ ID NO:225:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

CGGCCGAGTG TCTGGAGCAA

20

(2) INFORMATION FOR SEQ ID NO:226:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

GGATGAAGCG GAGTCTGGA

19

(2) INFORMATION FOR SEQ ID NO:227:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

ATGGATCCGT CGTCGAGCTG CTCAGGTCT

29

(2) INFORMATION FOR SEQ ID NO:228:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:

ATCAGCTGAG CACGCTGAAC AGTGCCTTC

29

(2) INFORMATION FOR SEQ ID NO:229:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:229:

GTCTCCGTGA CATAAAAGAA AGAC

24

(2) INFORMATION FOR SEQ ID NO:230:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:230:

GCCAAGTTCC TGCACTGGCT

20

(2) INFORMATION FOR SEQ ID NO:231:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:231:

GCCTGTTCTT TTGAAACGTG GTCT

24

(2) INFORMATION FOR SEQ ID NO:232:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /mod\_base= OTHER  
/note= "N = guanosine substituted by two  
biotin groups"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

NCCTGTTCTT TTGAAACGTG GTCT

24

(2) INFORMATION FOR SEQ ID NO:233:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

GTCAAGATGC CTGAGATAGA AC

22

(2) INFORMATION FOR SEQ ID NO:234:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

TGCTTAGCTT GTGGGGGTGT CA

22

(2) INFORMATION FOR SEQ ID NO:235:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:235:  
GCTGCGTCCT GCTGCGCACG T 21

(2) INFORMATION FOR SEQ ID NO:236:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:  
CAGCGGGGAG CGCGCGGCAT C 21

(2) INFORMATION FOR SEQ ID NO:237:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:  
TGGGCCACCA GCGCGCGGAA A 21

(2) INFORMATION FOR SEQ ID NO:238:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:  
CGGCCGCAGC CCGTCAGGCT TGGGG 25

(2) INFORMATION FOR SEQ ID NO:239:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:

CCGACAGCTC CCGCAGCTGC ACCC

24

(2) INFORMATION FOR SEQ ID NO:240:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:240:

CGTACACACT CATCAGCCAG TGCAGGAAGT TGGC

34

(2) INFORMATION FOR SEQ ID NO:241:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:241:

CGCGCCCGCT CGTAGTTGAG CACGCTGAAC AGTGCCTTC

39

(2) INFORMATION FOR SEQ ID NO:242:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:242:

GCGGAGTCTG GACGTCAGCA GGGCGGGCCT GGCTTCCCG

39

(2) INFORMATION FOR SEQ ID NO:243:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:243:

ATTGACCCA CAGGGACCCC CATCCAG

27

(2) INFORMATION FOR SEQ ID NO:244:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244:

ATGACCGCCC TCCTCGTGAG

20

(2) INFORMATION FOR SEQ ID NO:245:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:

GCCACCCCCG CGATGCC

17

(2) INFORMATION FOR SEQ ID NO:246:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:

AGCCCTGGCC CCGGCCA

17

(2) INFORMATION FOR SEQ ID NO:247:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

TCCCACGTGC GCAGCAG

17

(2) INFORMATION FOR SEQ ID NO:248:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

AGCAGGACGC AGCGCTG

17

(2) INFORMATION FOR SEQ ID NO:249:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

CGCGGTAGTG GCTGCGCAGC AGGGAGCGCA CGGC

34

(2) INFORMATION FOR SEQ ID NO:250:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

CCAGGGCTTC CCACGTGCGC AGCAGGACGC AGCGC

35



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:  
CACTGCTGGC CTCATTGAGG G

21

(2) INFORMATION FOR SEQ ID NO:213:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:  
GCAGCCATAC TCAGGGACAC

20

(2) INFORMATION FOR SEQ ID NO:214:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:  
CCATCCTCTC CACGCTGCTC

20

(2) INFORMATION FOR SEQ ID NO:215:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 21 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:  
GCGATGACCT CCGTGAGCCT G

21

(2) INFORMATION FOR SEQ ID NO:216:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 19 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:

CCCAGGACAG GCTCACGGA

19

(2) INFORMATION FOR SEQ ID NO:217:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:

CCTCTTCAAG TGCTGTCTGA TTCC

24

(2) INFORMATION FOR SEQ ID NO:218:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:

CAGCTCGACG ACGTACACAC TCATC

25

(2) INFORMATION FOR SEQ ID NO:219:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

CTGACGTCCA GACTCCGCTT CAT

23

(2) INFORMATION FOR SEQ ID NO:220:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

GACCTGAGCA GCTCGACGAC GTACACACTC ATC

33

(2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

GTCGTCGAGC TGCTCAGGTC

20

(2) INFORMATION FOR SEQ ID NO:222:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

AGCACGCTGA ACAGTGCCTT

20

(2) INFORMATION FOR SEQ ID NO:223:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

GACCTGAGCA GCTCGACGAC

20

(2) INFORMATION FOR SEQ ID NO:224:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

AAGGCACTGT TCAGCGTGCT

20

(2) INFORMATION FOR SEQ ID NO:225:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

CGGCCGAGTG TCTGGAGCAA

20

(2) INFORMATION FOR SEQ ID NO:226:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

GGATGAAGCG GAGTCTGGA

19

(2) INFORMATION FOR SEQ ID NO:227:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

ATGGATCCGT CGTCGAGCTG CTCAGGTCT

29

(2) INFORMATION FOR SEQ ID NO:228:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:

ATCAGCTGAG CACGCTGAAC AGTGCCTTC

29

(2) INFORMATION FOR SEQ ID NO:229:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:229:

GTCTCCGTGA CATAAAAGAA AGAC

24

(2) INFORMATION FOR SEQ ID NO:230:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:230:

GCCAAGTTCC TGCACTGGCT

20

(2) INFORMATION FOR SEQ ID NO:231:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:231:

GCCTGTTCTT TTGAAACGTG GTCT

24

(2) INFORMATION FOR SEQ ID NO:232:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified\_base  
 (B) LOCATION: 1  
 (D) OTHER INFORMATION: /mod\_base= OTHER  
 /note= "N = guanosine substituted by two  
 biotin groups"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

NCCTGTTCTT TTGAAACGTG GTCT

24

(2) INFORMATION FOR SEQ ID NO:233:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

GTCAAGATGC CTGAGATAGA AC

22

(2) INFORMATION FOR SEQ ID NO:234:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

TGCTTAGCTT GTGGGGGTGT CA

22

(2) INFORMATION FOR SEQ ID NO:235:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:235:  
GCTGCGTCCT GCTGCGCACG T 21

(2) INFORMATION FOR SEQ ID NO:236:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:  
CAGCGGGGAG CGCGCGGCAT C 21

(2) INFORMATION FOR SEQ ID NO:237:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:  
TGGGCCACCA GCGCGCGGAA A 21

(2) INFORMATION FOR SEQ ID NO:238:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:  
CGGCCGCGAGC CCGTCAGGCT TGGGG 25

(2) INFORMATION FOR SEQ ID NO:239:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:

CCGACAGCTC CCGCAGCTGC ACCC

24

(2) INFORMATION FOR SEQ ID NO:240:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:240:

CGTACACACT CATCAGCCAG TGCAGGAACT TGGC

34

(2) INFORMATION FOR SEQ ID NO:241:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:241:

CGCGCCCGCT CGTAGTTGAG CACGCTGAAC AGTGCCTTC

39

(2) INFORMATION FOR SEQ ID NO:242:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:242:

GCGGAGTCTG GACGTCAGCA GGGCGGGCCT GGCTTCCCG

39

(2) INFORMATION FOR SEQ ID NO:243:

- (i) SEQUENCE CHARACTERISTICS:



- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:243:

ATTTGACCCA CAGGGACCCC CATCCAG

27

(2) INFORMATION FOR SEQ ID NO:244:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244:

ATGACCGCCC TCCTCGTGAG

20

(2) INFORMATION FOR SEQ ID NO:245:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:

GCCACCCCCG CGATGCC

17

(2) INFORMATION FOR SEQ ID NO:246:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:

AGCCCTGGCC CCGGCCA

17

(2) INFORMATION FOR SEQ ID NO:247:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

TCCCACGTGC GCAGCAG

17

(2) INFORMATION FOR SEQ ID NO:248:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

AGCAGGACGC AGCGCTG

17

(2) INFORMATION FOR SEQ ID NO:249:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

CGCGGTAGTG GCTGCGCAGC AGGGAGCGCA CGGC

34

(2) INFORMATION FOR SEQ ID NO:250:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

CCAGGGCTTC CCACGTGCGC AGCAGGACGC AGCGC

35

(2) INFORMATION FOR SEQ ID NO:251:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 62 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:251:

CTAGTCTAGA TCRCTAGCGT AATCTGGAAC ATCGTATGGG TRTCCAGGAT GGTCTTGAAG 60

TC 62

(2) INFORMATION FOR SEQ ID NO:252:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:252:

TACCATGGGC TACCCATACG ACGTTCCAGA TTACGCTCA 39

(2) INFORMATION FOR SEQ ID NO:253:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:253:

TATGAGCGTA ATCTGGAACG TCGTATGGGT AGCCCATGG 39

(2) INFORMATION FOR SEQ ID NO:254:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 45 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:254:

GTGTACGTCG TCGAGCTCCT CAGGTCTGCC TTTTATGTCA CGGAG

45

(2) INFORMATION FOR SEQ ID NO:255:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:255:

GTGTACGTCG TCGAGCTCCT CAGGTCTTTC GCTTATGTCA CGGAGACC

48

(2) INFORMATION FOR SEQ ID NO:256:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

CCTCAGGTCT TTCTTTGCTG TCACGGAGAC AACGTTTCAA AAGAACAG

48

(2) INFORMATION FOR SEQ ID NO:257:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257:

GGTCTTTCTT TTATGTCGCG GAGACAACGT TTCAAAGAA CAG

43

(2) INFORMATION FOR SEQ ID NO:258:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:258:

CTTCTCTTTTA TGTCACGGCG ACAACGTTTC AAAAGAACA

39

(2) INFORMATION FOR SEQ ID NO:259:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:259:

ATGAGTGTGT ACGTCGTCGA GCTCCTCAGG TCTACCACGC AAAAGAACAG GCTCTTTTTC

60

(2) INFORMATION FOR SEQ ID NO:260:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:260:

GGCTGATGAG TGTGTACGTC GTCGA

25

(2) INFORMATION FOR SEQ ID NO:261:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:261:

ACGTGGTCTC CGTGACATAA AAGAA

25

(2) INFORMATION FOR SEQ ID NO:262:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:262:  
AGGTCTTTCT TTTATGTCAC GGA

23

(2) INFORMATION FOR SEQ ID NO:263:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:263:  
CACAGACCCC CGTCGCCTGG TC

22

(2) INFORMATION FOR SEQ ID NO:264:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:264:  
CGGAGTCTGG ACGTCAGCAG GGC

23

(2) INFORMATION FOR SEQ ID NO:265:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265:  
CGCGGATCCG TAACTAAAAT GCCGCGCGCT CCCCCTGC

39

(2) INFORMATION FOR SEQ ID NO:266:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:266:

CCGGAATTCG TTAGTTACTT ACAAAGAGGT GGCTTCTTCG GC

42

(2) INFORMATION FOR SEQ ID NO:267:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:267:

CGCGGATCCG TAACTAAAGC CACCTCTTTG GAGGGTGCG

39

(2) INFORMATION FOR SEQ ID NO:268:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:268:

CCGGAATTCG TTAGTTACTT AAGACCTGAG CAGCTCGACG AC

42

(2) INFORMATION FOR SEQ ID NO:269:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:269:

CGCGGATCCG TAACTAAAAT GAGTGTGTAC GTCGTCGAG

39

(2) INFORMATION FOR SEQ ID NO:270:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 40 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:270:  
 CCGGAATTCG TTAGTTACTT AGATCCCCTG GCACTGGACG 40

(2) INFORMATION FOR SEQ ID NO:271:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:271:  
 CGCGGATCCG TAACTAAAT CCCGCAGGGC TCCATCCTC 39

(2) INFORMATION FOR SEQ ID NO:272:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:272:  
 CCGGAATTCG TTAGTTACTT AGTCCAGGAT GGTCTTGAAG TC 42

(2) INFORMATION FOR SEQ ID NO:273:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:273:  
 GGCATCGCGG GGGTGGCCGG G 21

(2) INFORMATION FOR SEQ ID NO:274:

(i) SEQUENCE CHARACTERISTICS:



- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:274:

GGACACCTGG CGGAAGGAGG G

21

- (2) INFORMATION FOR SEQ ID NO:275:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:275:

GCGTGCCAGC AGGTGAACCA G

21

- (2) INFORMATION FOR SEQ ID NO:276:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:276:

CTCAGGGGCA GCGCCACGCC T

21

- (2) INFORMATION FOR SEQ ID NO:277:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:277:

AGGTGGCTTC TTCGGCGGGT C

21

(2) INFORMATION FOR SEQ ID NO:278:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:278:

GGACAAGGCG TGTCCAGGG A

21

(2) INFORMATION FOR SEQ ID NO:279:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:279:

GCTGGGGTGA CCGCAGCTCG C

21

(2) INFORMATION FOR SEQ ID NO:280:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:280:

GATGAACTTC TTGGTGTTCC T

21

(2) INFORMATION FOR SEQ ID NO:281:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:281:

GTGCGCCAGG CCCTGTGGAT A

21

(2) INFORMATION FOR SEQ ID NO:282:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:282:

GCCCATGGGC GGCCTTCTGG A

21

(2) INFORMATION FOR SEQ ID NO:283:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:283:

GAGGCCACTG CTGGCCTCAT T

21

(2) INFORMATION FOR SEQ ID NO:284:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:284:

GGGTGAGGTG AGGTGTCACC A

21

(2) INFORMATION FOR SEQ ID NO:285:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:285:

GCTGCAGCAC ACATGCGTGA AACCTGTACG C

31

- (2) INFORMATION FOR SEQ ID NO:286:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:286:

GACGCGCAGG AAAAATGTGG G

21

- (2) INFORMATION FOR SEQ ID NO:287:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:287:

CCGAGCGCCA GCCTGTGGGG A

21

- (2) INFORMATION FOR SEQ ID NO:288:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:288:

CAGCGGGGAG CGCGCGGCAT C

21

(2) INFORMATION FOR SEQ ID NO:289:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "phosphorothioate"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:289:

CAGCACCTCG CGGTAGTGGC T

21

(2) INFORMATION FOR SEQ ID NO:290:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:290:

TCAAGCCAAA CCTGAATCTG AG

22

(2) INFORMATION FOR SEQ ID NO:291:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:291:

CCCGAGTGAA TCTTTCTACG C

21

(2) INFORMATION FOR SEQ ID NO:292:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:292:

GTCTCTGGCA GTTCCTCAT CCC

23

(2) INFORMATION FOR SEQ ID NO:293:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:293:

TTTAGGCATC CTCCAAGCA CA

22

(2) INFORMATION FOR SEQ ID NO:294:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:294:

TTAGGGTTAG

10

(2) INFORMATION FOR SEQ ID NO:295:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:295:

TTAGGGTTAG GGTTAGGG

18

(2) INFORMATION FOR SEQ ID NO:296:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

GTTAGGGTTA GGGTTAGG

18

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: repeat\_unit  
(B) LOCATION: 1..6  
(D) OTHER INFORMATION: /note= "sequence (CCCTAA)-n, where n is at least 1, or at least 3, or at least 10 or more"

CCCTAACCCT AACCTAACC CTAACCCTAA CCCTAACCCT AACCTAACC CTAACCCTAA

60

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..30  
(D) OTHER INFORMATION: /note= "non-telomeric nucleotide  
sequence, (N)-n, where n is 8-20,  
or 6-30"

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN TTAG

34

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..30
- (D) OTHER INFORMATION: /note= "non-telomeric nucleotide sequence, (N)-n, where n is 8-20, or 6-30"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:299:

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN AGGG

34

(2) INFORMATION FOR SEQ ID NO:300:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..30
- (D) OTHER INFORMATION: /note= "non-telomeric nucleotide sequence, (N)-n, where n is 8-20, or 6-30"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:300:

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN TTAGGGTTAG

40

(2) INFORMATION FOR SEQ ID NO:301:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..30
- (D) OTHER INFORMATION: /note= "non-telomeric nucleotide sequence, (N)-n, where n is 8-20, or 6-30"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:301:

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN TTAGGGTTAG GGTTAG

46



(2) INFORMATION FOR SEQ ID NO:302:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 52 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: 1..30  
(D) OTHER INFORMATION: /note= "non-telomeric nucleotide  
sequence, (N)-n, where n is 8-20,  
or 6-30"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:302:

NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN TTAGGGTTAG GGTTAGGGTT AG

52

(2) INFORMATION FOR SEQ ID NO:303:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 58 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION: 1..30  
(D) OTHER INFORMATION: /note= "non-telomeric nucleotide  
sequence, (N)-n, where n is 8-20,  
or 6-30"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:303:

NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN TTAGGGTTAG GGTTAGGGTT AGGGTTAG

58

(2) INFORMATION FOR SEQ ID NO:304:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:304:

TAGGGATTAG

10

(2) INFORMATION FOR SEQ ID NO:305:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 16
- (D) OTHER INFORMATION: /mod\_base= OTHER  
/note= "N = 3'-deoxyguanosine"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:305:

TTAGGGTTAG GGTAN

16

(2) INFORMATION FOR SEQ ID NO:306:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: repeat\_unit
- (B) LOCATION: 1..6
- (D) OTHER INFORMATION: /note= "sequence (TTAGGG)-n, where n  
is 1-10, or typically 3-5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:306:

TTAGGGTTAG GGTAGGGTT AGGGTTAGGG TTAGGGTTAG GGTAGGGTT AGGGTTAGGG

60

(2) INFORMATION FOR SEQ ID NO:307:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:307:

GGCACTGGAC GTAGGACGTG

20

(2) INFORMATION FOR SEQ ID NO:308:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:308:

CGGAAGAGTG TCTGGAGCAA

20

(2) INFORMATION FOR SEQ ID NO:309:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:309:

CTCAGACACC ATGGGGAAGG TGA

23

(2) INFORMATION FOR SEQ ID NO:310:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:310:

ATGATCTTGA GGCTGTTGTC ATA

23

(2) INFORMATION FOR SEQ ID NO:311:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:311:

26

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:312:

26

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 170 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:313:

CCCCCGCCG	CCCCCTCCTT	CCGCCAGGTG	GGCTTCCCCG	GGGTCGGCGT	CCGGCTGGGG	60
TTGAGGGCGG	CCGGGGGGAA	CCAGCGACAT	GCGGAGAGCA	GCGCAGGCGA	CTCAGGGCGC	120
TTCCCCCGCA	GGTGTCTGTC	CTGAAGGAGC	TGGTGGCCCCG	AGTGCTGCAG		170

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1285 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:314:

Met	Ser	Asp	Lys	Ile	Ile	His	Leu	Thr	Asp	Asp	Ser	Phe	Asp	Thr	Asp
1				5					10					15	
Val	Leu	Lys	Ala	Asp	Gly	Ala	Ile	Leu	Val	Asp	Phe	Trp	Ala	His	Trp
			20					25					30		
Cys	Gly	Pro	Cys	Lys	Met	Ile	Ala	Pro	Ile	Leu	Asp	Glu	Ile	Ala	Asp
		35					40					45			
Glu	Tyr	Gln	Gly	Lys	Leu	Thr	Val	Ala	Lys	Leu	Arg	Ile	Asp	His	Asn

50					55					60					
Pro	Gly	Thr	Ala	Pro	Lys	Tyr	Gly	Ile	Arg	Gly	Ile	Pro	Thr	Leu	Leu
65					70					75					80
Leu	Phe	Lys	Asn	Gly	Glu	Val	Ala	Ala	Thr	Lys	Val	Gly	Ala	Leu	Ser
				85					90					95	
Lys	Gly	Gln	Leu	Lys	Glu	Phe	Leu	Asp	Ala	Asn	Leu	Ala	Gly	Ser	Gly
			100					105					110		
Ser	Gly	Asp	Asp	Asp	Asp	Lys	Val	Pro	Met	His	Glu	Leu	Glu	Ile	Phe
		115					120					125			
Glu	Phe	Ala	Ala	Ala	Ser	Thr	Gln	Arg	Cys	Val	Leu	Leu	Arg	Thr	Trp
	130					135					140				
Glu	Ala	Leu	Ala	Pro	Ala	Thr	Pro	Ala	Met	Pro	Arg	Ala	Pro	Arg	Cys
145					150					155					160
Arg	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser	His	Tyr	Arg	Glu	Val	Leu	Pro
				165					170					175	
Leu	Ala	Thr	Phe	Val	Arg	Arg	Leu	Gly	Pro	Gln	Gly	Trp	Arg	Leu	Val
			180					185					190		
Gln	Arg	Gly	Asp	Pro	Ala	Ala	Phe	Arg	Ala	Leu	Val	Ala	Gln	Cys	Leu
		195					200					205			
Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro	Pro	Pro	Ala	Ala	Pro	Ser	Phe
	210					215					220				
Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Leu	Val	Ala	Arg	Val	Leu	Gln	Arg
225					230					235					240
Leu	Cys	Glu	Arg	Gly	Ala	Lys	Asn	Val	Leu	Ala	Phe	Gly	Phe	Ala	Leu
				245					250					255	
Leu	Asp	Gly	Ala	Arg	Gly	Gly	Pro	Pro	Glu	Ala	Phe	Thr	Thr	Ser	Val
		260					265						270		
Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Thr	Asp	Ala	Leu	Arg	Gly	Ser	Gly
		275					280					285			
Ala	Trp	Gly	Leu	Leu	Leu	Arg	Arg	Val	Gly	Asp	Asp	Val	Leu	Val	His
	290					295					300				
Leu	Leu	Ala	Arg	Cys	Ala	Leu	Phe	Val	Leu	Val	Ala	Pro	Ser	Cys	Ala
305					310					315					320
Tyr	Gln	Val	Cys	Gly	Pro	Pro	Leu	Tyr	Gln	Leu	Gly	Ala	Ala	Thr	Gln
				325					330					335	
Ala	Arg	Pro	Pro	Pro	His	Ala	Ser	Gly	Pro	Arg	Arg	Arg	Leu	Gly	Cys
			340					345					350		
Glu	Arg	Ala	Trp	Asn	His	Ser	Val	Arg	Glu	Ala	Gly	Val	Pro	Leu	Gly
		355					360					365			
Leu	Pro	Ala	Pro	Gly	Ala	Arg	Arg	Arg	Gly	Gly	Ser	Ala	Ser	Arg	Ser
	370					375					380				

Leu Pro Leu Pro Lys Arg Pro Arg Arg Gly Ala Ala Pro Glu Pro Glu  
 385 390 395 400  
 Arg Thr Pro Val Gly Gln Gly Ser Trp Ala His Pro Gly Arg Thr Arg  
 405 410 415  
 Gly Pro Ser Asp Arg Gly Phe Cys Val Val Ser Pro Ala Arg Pro Ala  
 420 425 430  
 Glu Glu Ala Thr Ser Leu Glu Gly Ala Leu Ser Gly Thr Arg His Ser  
 435 440 445  
 His Pro Ser Val Gly Arg Gln His His Ala Gly Pro Pro Ser Thr Ser  
 450 455 460  
 Arg Pro Pro Arg Pro Trp Asp Thr Pro Cys Pro Pro Val Tyr Ala Glu  
 465 470 475 480  
 Thr Lys His Phe Leu Tyr Ser Ser Gly Asp Lys Glu Gln Leu Arg Pro  
 485 490 495  
 Ser Phe Leu Leu Ser Ser Leu Arg Pro Ser Leu Thr Gly Ala Arg Arg  
 500 505 510  
 Leu Val Glu Thr Ile Phe Leu Gly Ser Arg Pro Trp Met Pro Gly Thr  
 515 520 525  
 Pro Arg Arg Leu Pro Arg Leu Pro Gln Arg Tyr Trp Gln Met Arg Pro  
 530 535 540  
 Leu Phe Leu Glu Leu Leu Gly Asn His Ala Gln Cys Pro Tyr Gly Val  
 545 550 555 560  
 Leu Leu Lys Thr His Cys Pro Leu Arg Ala Ala Val Thr Pro Ala Ala  
 565 570 575  
 Gly Val Cys Ala Arg Glu Lys Pro Gln Gly Ser Val Ala Ala Pro Glu  
 580 585 590  
 Glu Glu Asp Thr Asp Pro Arg Arg Leu Val Gln Leu Leu Arg Gln His  
 595 600 605  
 Ser Ser Pro Trp Gln Val Tyr Gly Phe Val Arg Ala Cys Leu Arg Arg  
 610 615 620  
 Leu Val Pro Pro Gly Leu Trp Gly Ser Arg His Asn Glu Arg Arg Phe  
 625 630 635 640  
 Leu Arg Asn Thr Lys Lys Phe Ile Ser Leu Gly Lys His Ala Lys Leu  
 645 650 655  
 Ser Leu Gln Glu Leu Thr Trp Lys Met Ser Val Arg Asp Cys Ala Trp  
 660 665 670  
 Leu Arg Arg Ser Pro Gly Val Gly Cys Val Pro Ala Ala Glu His Arg  
 675 680 685  
 Leu Arg Glu Glu Ile Leu Ala Lys Phe Leu His Trp Leu Met Ser Val  
 690 695 700  
 Tyr Val Val Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr  
 705 710 715 720

Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val Trp Ser Lys  
725 730 735  
Leu Gln Ser Ile Gly Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg  
740 745 750  
Glu Leu Ser Glu Ala Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala  
755 760 765  
Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg  
770 775 780  
Pro Ile Val Asn Met Asp Tyr Val Val Gly Ala Arg Thr Phe Arg Arg  
785 790 795 800  
Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg Val Lys Ala Leu Phe Ser  
805 810 815  
Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser  
820 825 830  
Val Leu Gly Leu Asp Asp Ile His Arg Ala Trp Arg Thr Phe Val Leu  
835 840 845  
Arg Val Arg Ala Gln Asp Pro Pro Pro Glu Leu Tyr Phe Val Lys Val  
850 855 860  
Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu  
865 870 875 880  
Val Ile Ala Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg  
885 890 895  
Tyr Ala Val Val Gln Lys Ala Ala His Gly His Val Arg Lys Ala Phe  
900 905 910  
Lys Ser His Val Ser Thr Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln  
915 920 925  
Phe Val Ala His Leu Gln Glu Thr Ser Pro Leu Arg Asp Ala Val Val  
930 935 940  
Ile Glu Gln Ser Ser Ser Leu Asn Glu Ala Ser Ser Gly Leu Phe Asp  
945 950 955 960  
Val Phe Leu Arg Phe Met Cys His His Ala Val Arg Ile Arg Gly Lys  
965 970 975  
Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln Gly Ser Ile Leu Ser Thr  
980 985 990  
Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met Glu Asn Lys Leu Phe Ala  
995 1000 1005  
Gly Ile Arg Arg Asp Gly Leu Leu Leu Arg Leu Val Asp Asp Phe Leu  
1010 1015 1020  
Leu Val Thr Pro His Leu Thr His Ala Lys Thr Phe Leu Arg Thr Leu  
1025 1030 1035 1040  
Val Arg Gly Val Pro Glu Tyr Gly Cys Val Val Asn Leu Arg Lys Thr  
1045 1050 1055

Val Val Asn Phe Pro Val Glu Asp Glu Ala Leu Gly Gly Thr Ala Phe  
 1060 1065 1070  
 Val Gln Met Pro Ala His Gly Leu Phe Pro Trp Cys Gly Leu Leu Leu  
 1075 1080 1085  
 Asp Thr Arg Thr Leu Glu Val Gln Ser Asp Tyr Ser Ser Tyr Ala Arg  
 1090 1095 1100  
 Thr Ser Ile Arg Ala Ser Leu Thr Phe Asn Arg Gly Phe Lys Ala Gly  
 1105 1110 1115 1120  
 Arg Asn Met Arg Arg Lys Leu Phe Gly Val Leu Arg Leu Lys Cys His  
 1125 1130 1135  
 Ser Leu Phe Leu Asp Leu Gln Val Asn Ser Leu Gln Thr Val Cys Thr  
 1140 1145 1150  
 Asn Ile Tyr Lys Ile Leu Leu Leu Gln Ala Tyr Arg Phe His Ala Cys  
 1155 1160 1165  
 Val Leu Gln Leu Pro Phe His Gln Gln Val Trp Lys Asn Pro Thr Phe  
 1170 1175 1180  
 Phe Leu Arg Val Ile Ser Asp Thr Ala Ser Leu Cys Tyr Ser Ile Leu  
 1185 1190 1195 1200  
 Lys Ala Lys Asn Ala Gly Met Ser Leu Gly Ala Lys Gly Ala Ala Gly  
 1205 1210 1215  
 Pro Leu Pro Ser Glu Ala Val Gln Trp Leu Cys His Gln Ala Phe Leu  
 1220 1225 1230  
 Leu Lys Leu Thr Arg His Arg Val Thr Tyr Val Pro Leu Leu Gly Ser  
 1235 1240 1245  
 Leu Arg Thr Ala Gln Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr  
 1250 1255 1260  
 Leu Thr Ala Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe  
 1265 1270 1275 1280  
 Lys Thr Ile Leu Asp  
 1285

(2) INFORMATION FOR SEQ ID NO:315:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:315:

Gly Ser Val Thr Lys  
 1 5



(2) INFORMATION FOR SEQ ID NO:316:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 538 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:316:

Met	Ser	Pro	Ile	Leu	Gly	Tyr	Trp	Lys	Ile	Lys	Gly	Leu	Val	Gln	Pro	
1				5					10					15		
Thr	Arg	Leu	Leu	Leu	Glu	Tyr	Leu	Glu	Glu	Lys	Tyr	Glu	Glu	His	Leu	
			20					25					30			
Tyr	Glu	Arg	Asp	Glu	Gly	Asp	Lys	Trp	Arg	Asn	Lys	Lys	Phe	Glu	Leu	
		35					40					45				
Gly	Leu	Glu	Phe	Pro	Asn	Leu	Pro	Tyr	Tyr	Ile	Asp	Gly	Asp	Val	Lys	
	50					55					60					
Leu	Thr	Gln	Ser	Met	Ala	Ile	Ile	Arg	Tyr	Ile	Ala	Asp	Lys	His	Asn	
65					70					75					80	
Met	Leu	Gly	Gly	Cys	Pro	Lys	Glu	Arg	Ala	Glu	Ile	Ser	Met	Leu	Glu	
				85					90					95		
Gly	Ala	Val	Leu	Asp	Ile	Arg	Tyr	Gly	Val	Ser	Arg	Ile	Ala	Tyr	Ser	
			100					105					110			
Lys	Asp	Phe	Glu	Thr	Leu	Lys	Val	Asp	Phe	Leu	Ser	Lys	Leu	Pro	Glu	
		115					120					125				
Met	Leu	Lys	Met	Phe	Glu	Asp	Arg	Leu	Cys	His	Lys	Thr	Tyr	Leu	Asn	
	130					135					140					
Gly	Asp	His	Val	Thr	His	Pro	Asp	Phe	Met	Leu	Tyr	Asp	Ala	Leu	Asp	
145					150					155					160	
Val	Val	Leu	Tyr	Met	Asp	Pro	Met	Cys	Leu	Asp	Ala	Phe	Pro	Lys	Leu	
				165					170					175		
Val	Cys	Phe	Lys	Lys	Arg	Ile	Glu	Ala	Ile	Pro	Gln	Ile	Asp	Lys	Tyr	
			180					185					190			
Leu	Lys	Ser	Ser	Lys	Tyr	Ile	Ala	Trp	Pro	Leu	Gln	Gly	Trp	Gln	Ala	
		195					200					205				
Thr	Phe	Gly	Gly	Gly	Asp	His	Pro	Pro	Lys	Ser	Asp	Leu	Val	Pro	Arg	
	210					215					220					
Gly	Ser	Arg	Arg	Ala	Ser	Val	Gly	Ser	Val	Thr	Lys	Ile	Pro	Gln	Gly	
225					230					235					240	
Ser	Ile	Leu	Ser	Thr	Leu	Leu	Cys	Ser	Leu	Cys	Tyr	Gly	Asp	Met	Glu	
				245					250					255		
Asn	Lys	Leu	Phe	Ala	Gly	Ile	Arg	Arg	Asp	Gly	Leu	Leu	Leu	Arg	Leu	

260										265										270																									
Val	Asp	Asp	Phe	Leu	Leu	Val	Thr	Pro	His	Leu	Thr	His	Ala	Lys	Thr																														
		275					280					285																																	
Phe	Leu	Arg	Thr	Leu	Val	Arg	Gly	Val	Pro	Glu	Tyr	Gly	Cys	Val	Val																														
	290					295					300																																		
Asn	Leu	Arg	Lys	Thr	Val	Val	Asn	Phe	Pro	Val	Glu	Asp	Glu	Ala	Leu																														
	305				310					315					320																														
Gly	Gly	Thr	Ala	Phe	Val	Gln	Met	Pro	Ala	His	Gly	Leu	Phe	Pro	Trp																														
				325					330					335																															
Cys	Gly	Leu	Leu	Leu	Asp	Thr	Arg	Thr	Leu	Glu	Val	Gln	Ser	Asp	Tyr																														
			340					345					350																																
Ser	Ser	Tyr	Ala	Arg	Thr	Ser	Ile	Arg	Ala	Ser	Val	Thr	Phe	Asn	Arg																														
		355					360					365																																	
Gly	Phe	Lys	Ala	Gly	Arg	Asn	Met	Arg	Arg	Lys	Leu	Phe	Gly	Val	Leu																														
	370					375					380																																		
Arg	Leu	Lys	Cys	His	Ser	Leu	Phe	Leu	Asp	Leu	Gln	Val	Asn	Ser	Leu																														
	385				390				395					400																															
Gln	Thr	Val	Cys	Thr	Asn	Ile	Tyr	Lys	Ile	Leu	Leu	Leu	Gln	Ala	Tyr																														
				405				410						415																															
Arg	Phe	His	Ala	Cys	Val	Leu	Gln	Leu	Pro	Phe	His	Gln	Gln	Val	Trp																														
			420					425					430																																
Lys	Asn	Pro	Thr	Phe	Phe	Leu	Arg	Val	Ile	Ser	Asp	Thr	Ala	Ser	Leu																														
		435					440					445																																	
Cys	Tyr	Ser	Ile	Leu	Lys	Ala	Lys	Asn	Ala	Gly	Met	Ser	Leu	Gly	Ala																														
	450					455				460																																			
Lys	Gly	Ala	Ala	Gly	Pro	Leu	Pro	Ser	Glu	Ala	Val	Gln	Trp	Leu	Cys																														
	465				470				475					480																															
His	Gln	Ala	Phe	Leu	Leu	Lys	Leu	Thr	Arg	His	Arg	Val	Thr	Tyr	Val																														
				485				490					495																																
Pro	Leu	Leu	Gly	Ser	Leu	Arg	Thr	Ala	Gln	Thr	Gln	Leu	Ser	Arg	Lys																														
			500					505					510																																
Leu	Pro	Gly	Thr	Thr	Leu	Thr	Ala	Leu	Glu	Ala	Ala	Ala	Asn	Pro	Ala																														
		515					520					525																																	
Leu	Pro	Ser	Asp	Phe	Lys	Thr	Ile	Leu	Asp																																				
	530					535																																							

(2) INFORMATION FOR SEQ ID NO:317:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 530 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:317:

Met	Ser	Pro	Ile	Leu	Gly	Tyr	Trp	Lys	Ile	Lys	Gly	Leu	Val	Gln	Pro	
1				5					10					15		
Thr	Arg	Leu	Leu	Leu	Glu	Tyr	Leu	Glu	Glu	Lys	Tyr	Glu	Glu	His	Leu	
			20					25					30			
Tyr	Glu	Arg	Asp	Glu	Gly	Asp	Lys	Trp	Arg	Asn	Lys	Lys	Phe	Glu	Leu	
		35					40					45				
Gly	Leu	Glu	Phe	Pro	Asn	Leu	Pro	Tyr	Tyr	Ile	Asp	Gly	Asp	Val	Lys	
	50					55					60					
Leu	Thr	Gln	Ser	Met	Ala	Ile	Ile	Arg	Tyr	Ile	Ala	Asp	Lys	His	Asn	
65					70					75					80	
Met	Leu	Gly	Gly	Cys	Pro	Lys	Glu	Arg	Ala	Glu	Ile	Ser	Met	Leu	Glu	
				85					90					95		
Gly	Ala	Val	Leu	Asp	Ile	Arg	Tyr	Gly	Val	Ser	Arg	Ile	Ala	Tyr	Ser	
			100					105					110			
Lys	Asp	Phe	Glu	Thr	Leu	Lys	Val	Asp	Phe	Leu	Ser	Lys	Leu	Pro	Glu	
		115					120					125				
Met	Leu	Lys	Met	Phe	Glu	Asp	Arg	Leu	Cys	His	Lys	Thr	Tyr	Leu	Asn	
	130					135					140					
Gly	Asp	His	Val	Thr	His	Pro	Asp	Phe	Met	Leu	Tyr	Asp	Ala	Leu	Asp	
145					150					155					160	
Val	Val	Leu	Tyr	Met	Asp	Pro	Met	Cys	Leu	Asp	Ala	Phe	Pro	Lys	Leu	
				165					170					175		
Val	Cys	Phe	Lys	Lys	Arg	Ile	Glu	Ala	Ile	Pro	Gln	Ile	Asp	Lys	Tyr	
			180					185					190			
Leu	Lys	Ser	Ser	Lys	Tyr	Ile	Ala	Trp	Pro	Leu	Gln	Gly	Trp	Gln	Ala	
		195					200					205				
Thr	Phe	Gly	Gly	Gly	Asp	His	Pro	Pro	Lys	Ser	Asp	Leu	Val	Pro	Arg	
	210					215					220					
Gly	Ser	Arg	Arg	Ala	Ser	Val	Gly	Ser	Val	His	His	His	His	His	His	
225					230					235					240	
His	His	Gly	Ser	Val	Thr	Lys	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu	
				245					250					255		
Arg	Ser	Phe	Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu	
			260					265					270			
Phe	Phe	Tyr	Arg	Pro	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile	
		275					280					285				
Arg	Gln	His	Leu	Lys	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser	Glu	Ala	Glu	
	290					295					300					

Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu  
 305 310 315 320  
 Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp  
 325 330 335  
 Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg  
 340 345 350  
 Leu Thr Ser Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg  
 355 360 365  
 Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp  
 370 375 380  
 Ile His Arg Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp  
 385 390 395 400  
 Pro Pro Pro Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr  
 405 410 415  
 Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile  
 420 425 430  
 Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys  
 435 440 445  
 Ala Ala His Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr  
 450 455 460  
 Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln  
 465 470 475 480  
 Glu Thr Ser Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser  
 485 490 495  
 Leu Asn Glu Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met  
 500 505 510  
 Cys His His Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln  
 515 520 525  
 Gly Ile  
 530

(2) INFORMATION FOR SEQ ID NO:318:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 515 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:318:

Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro  
 1 5 10 15

Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu  
 20 25 30  
 Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu  
 35 40 45  
 Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys  
 50 55 60  
 Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn  
 65 70 75 80  
 Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu  
 85 90 95  
 Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser  
 100 105 110  
 Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu  
 115 120 125  
 Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn  
 130 135 140  
 Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp  
 145 150 155 160  
 Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu  
 165 170 175  
 Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr  
 180 185 190  
 Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala  
 195 200 205  
 Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Val Pro Arg  
 210 215 220  
 Gly Ser Arg Arg Ala Ser Val Gly Ser Val Thr Lys Met Ser Val Tyr  
 225 230 235 240  
 Val Val Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr Phe  
 245 250 255  
 Gln Lys Asn Arg Leu Phe Phe Tyr Arg Pro Ser Val Trp Ser Lys Leu  
 260 265 270  
 Gln Ser Ile Gly Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg Glu  
 275 280 285  
 Leu Ser Glu Ala Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu  
 290 295 300  
 Leu Thr Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro  
 305 310 315 320  
 Ile Val Asn Met Asp Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu  
 325 330 335  
 Lys Arg Ala Glu Arg Leu Thr Ser Arg Lys Ala Leu Phe Ser Val Leu  
 340 345 350

Asn Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser Val Leu  
 355 360 365  
 Gly Leu Asp Asp Ile His Arg Ala Trp Arg Thr Phe Val Leu Arg Val  
 370 375 380  
 Arg Ala Gln Asp Pro Pro Pro Glu Tyr Phe Val Lys Val Asp Val Thr  
 385 390 395 400  
 Gly Ala Tyr Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala  
 405 410 415  
 Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val  
 420 425 430  
 Val Gln Lys Ala Ala His Gly Val Arg Lys Ala Phe Lys Ser His Val  
 435 440 445  
 Ser Thr Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His  
 450 455 460  
 Leu Gln Glu Thr Ser Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser  
 465 470 475 480  
 Ser Ser Leu Asn Glu Ala Ser Gly Leu Phe Asp Val Phe Leu Arg Phe  
 485 490 495  
 Met Cys His His Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys  
 500 505 510  
 Gln Gly Ile  
 515

(2) INFORMATION FOR SEQ ID NO:319:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 514 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:319:

Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro  
 1 5 10 15  
 Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu  
 20 25 30  
 Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu  
 35 40 45  
 Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys  
 50 55 60  
 Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn  
 65 70 75 80

Met	Leu	Gly	Gly	Cys	Pro	Lys	Glu	Arg	Ala	Glu	Ile	Ser	Met	Leu	Glu	85	90	95
Gly	Ala	Val	Leu	Asp	Ile	Arg	Tyr	Gly	Val	Ser	Arg	Ile	Ala	Tyr	Ser	100	105	110
Lys	Asp	Phe	Glu	Thr	Leu	Lys	Val	Asp	Phe	Leu	Ser	Lys	Leu	Pro	Glu	115	120	125
Met	Leu	Lys	Met	Phe	Glu	Asp	Arg	Leu	Cys	His	Lys	Thr	Tyr	Leu	Asn	130	135	140
Gly	Asp	His	Val	Thr	His	Pro	Asp	Phe	Met	Leu	Tyr	Asp	Ala	Leu	Asp	145	150	155
Val	Val	Leu	Tyr	Met	Asp	Pro	Met	Cys	Leu	Asp	Ala	Phe	Pro	Lys	Leu	165	170	175
Val	Cys	Phe	Lys	Lys	Arg	Ile	Glu	Ala	Ile	Pro	Gln	Ile	Asp	Lys	Tyr	180	185	190
Leu	Lys	Ser	Ser	Lys	Tyr	Ile	Ala	Trp	Pro	Leu	Gln	Gly	Trp	Gln	Ala	195	200	205
Thr	Phe	Gly	Gly	Gly	Asp	His	Pro	Pro	Lys	Ser	Asp	Leu	Val	Pro	Arg	210	215	220
Gly	Ser	Arg	Arg	Ala	Ser	Val	Gly	Ser	Val	Thr	Lys	Ala	Thr	Ser	Leu	225	230	235
Glu	Gly	Ala	Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg	245	250	255
Gln	His	His	Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp	260	265	270
Asp	Thr	Pro	Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr	275	280	285
Ser	Ser	Gly	Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser	290	295	300
Leu	Arg	Pro	Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe	305	310	315
Leu	Gly	Ser	Arg	Pro	Trp	Met	Pro	Gly	Thr	Pro	Arg	Arg	Leu	Pro	Arg	325	330	335
Leu	Pro	Gln	Arg	Tyr	Trp	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu	Leu	340	345	350
Gly	Asn	His	Ala	Gln	Cys	Pro	Tyr	Gly	Val	Leu	Leu	Lys	Thr	His	Cys	355	360	365
Pro	Leu	Arg	Ala	Ala	Val	Thr	Pro	Ala	Ala	Gly	Val	Cys	Ala	Arg	Glu	370	375	380
Lys	Pro	Gln	Gly	Ser	Val	Ala	Ala	Pro	Glu	Glu	Glu	Asp	Thr	Asp	Pro	385	390	395
Arg	Arg	Leu	Val	Gln	Leu	Leu	Arg	Gln	His	Ser	Ser	Pro	Trp	Gln	Val	405	410	415

Tyr Gly Phe Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu  
 420 425 430  
 Trp Gly Ser Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys  
 435 440 445  
 Phe Ile Ser Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr  
 450 455 460  
 Trp Lys Met Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly  
 465 470 475 480  
 Val Gly Cys Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu  
 485 490 495  
 Ala Lys Phe Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu  
 500 505 510  
 Arg Ser

(2) INFORMATION FOR SEQ ID NO:320:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 517 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:320:

Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro  
 1 5 10 15  
 Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu  
 20 25 30  
 Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu  
 35 40 45  
 Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys  
 50 55 60  
 Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn  
 65 70 75 80  
 Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu  
 85 90 95  
 Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser  
 100 105 110  
 Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu  
 115 120 125  
 Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn  
 130 135 140  
 Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp



145		150		155		160
Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu	165		170		175	
Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr	180		185		190	
Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala	195		200		205	
Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Val Pro Arg	210		215		220	
Gly Ser Arg Arg Ala Ser Val Gly Ser Val Thr Lys Met Pro Arg Ala	225		230		235	240
Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Ser His Tyr Arg Glu Val	245		250		255	
Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly Pro Gln Gly Trp Arg	260		265		270	
Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg Ala Leu Val Ala Gln	275		280		285	
Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro Pro Ala Ala Pro Ser	290		295		300	
Phe Arg Gln Val Ser Cys Leu Lys Glu Leu Val Ala Arg Val Leu Gln	305		310		315	320
Arg Leu Cys Glu Arg Gly Ala Lys Asn Val Leu Ala Phe Gly Phe Ala	325		330		335	
Leu Leu Asp Gly Ala Arg Gly Gly Pro Pro Glu Ala Thr Thr Ser Val	340		345		350	
Arg Ser Tyr Leu Pro Asn Thr Val Thr Asp Ala Leu Arg Gly Ser Gly	355		360		365	
Ala Trp Gly Leu Leu Leu Arg Arg Val Gly Asp Asp Val Leu Val His	370		375		380	
Leu Leu Ala Arg Cys Ala Leu Phe Val Leu Val Ala Pro Cys Ala Tyr	385		390		395	400
Gln Val Cys Gly Pro Pro Leu Tyr Gln Leu Gly Ala Ala Thr Gln Ala	405		410		415	
Arg Pro Pro Pro His Ala Ser Gly Pro Arg Arg Arg Leu Gly Cys Glu	420		425		430	
Arg Ala Trp Asn His Ser Val Arg Glu Ala Gly Val Pro Leu Gly Leu	435		440		445	
Pro Ala Pro Gly Ala Arg Arg Arg Gly Gly Ser Ala Ser Arg Ser Leu	450		455		460	
Pro Leu Pro Lys Arg Pro Arg Arg Gly Ala Ala Pro Glu Pro Glu Arg	465		470		475	480

Thr Pro Val Gly Gln Gly Ser Trp Ala His Pro Gly Arg Thr Arg Gly  
485 490 495

Pro Ser Asp Arg Gly Phe Cys Val Val Ser Pro Ala Arg Pro Ala Glu  
500 505 510

Glu Ala Thr Ser Leu  
515

(2) INFORMATION FOR SEQ ID NO:321:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:321:

CCGGCCACCC CCCATATGCC GCGCGCTCCC

30

(2) INFORMATION FOR SEQ ID NO:322:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:322:

Asn Ser Ala Val Asp  
1 5

(2) INFORMATION FOR SEQ ID NO:323:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1154 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:323:

Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser  
1 5 10 15

His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly  
20 25 30

Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg

35					40					45					
Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro
50						55					60				
Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Leu
65					70					75					80
Val	Ala	Arg	Val	Leu	Gln	Arg	Leu	Cys	Glu	Arg	Gly	Ala	Lys	Asn	Val
				85					90					95	
Leu	Ala	Phe	Gly	Phe	Ala	Leu	Leu	Asp	Gly	Ala	Arg	Gly	Gly	Pro	Pro
			100					105					110		
Glu	Ala	Phe	Thr	Thr	Ser	Val	Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Thr
			115				120					125			
Asp	Ala	Leu	Arg	Gly	Ser	Gly	Ala	Trp	Gly	Leu	Leu	Leu	Arg	Arg	Val
						135					140				
Gly	Asp	Asp	Val	Leu	Val	His	Leu	Leu	Ala	Arg	Cys	Ala	Leu	Phe	Val
145				150						155					160
Leu	Val	Ala	Pro	Ser	Cys	Ala	Tyr	Gln	Val	Cys	Gly	Pro	Pro	Leu	Tyr
				165					170					175	
Gln	Leu	Gly	Ala	Ala	Thr	Gln	Ala	Arg	Pro	Pro	Pro	Pro	His	Ala	Ser
			180					185						190	
Pro	Arg	Arg	Arg	Leu	Gly	Cys	Glu	Arg	Ala	Trp	Asn	His	Ser	Val	Arg
			195				200					205			
Glu	Ala	Gly	Val	Pro	Leu	Gly	Leu	Pro	Ala	Pro	Gly	Ala	Arg	Arg	Arg
			210			215					220				
Gly	Gly	Ser	Ala	Ser	Arg	Ser	Leu	Pro	Leu	Pro	Lys	Arg	Pro	Arg	Arg
225				230						235				240	
Gly	Ala	Ala	Pro	Glu	Pro	Glu	Arg	Thr	Pro	Val	Gly	Gln	Gly	Ser	Trp
				245					250					255	
Ala	His	Pro	Gly	Arg	Thr	Arg	Gly	Pro	Ser	Asp	Arg	Gly	Phe	Cys	Val
			260				265						270		
Val	Ser	Pro	Ala	Arg	Pro	Ala	Glu	Glu	Ala	Thr	Ser	Leu	Glu	Gly	Ala
			275				280					285			
Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg	Gln	His	His
			290			295					300				
Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp	Asp	Thr	Pro
305				310						315					320
Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr	Ser	Ser	Gly
				325					330					335	
Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser	Leu	Arg	Pro
			340					345					350		
Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe	Leu	Gly	Ser
			355				360					365			

Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln  
 370 375 380  
 Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His  
 385 390 395 400  
 Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg  
 405 410 415  
 Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln  
 420 425 430  
 Gly Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp Pro Arg Arg Leu  
 435 440 445  
 Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe  
 450 455 460  
 Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser  
 465 470 475 480  
 Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser  
 485 490 495  
 Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met  
 500 505 510  
 Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys  
 515 520 525  
 Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe  
 530 535 540  
 Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe  
 545 550 555 560  
 Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr  
 565 570 575  
 Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His  
 580 585 590  
 Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln  
 595 600 605  
 His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile  
 610 615 620  
 Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val  
 625 630 635 640  
 Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser  
 645 650 655  
 Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg  
 660 665 670  
 Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg  
 675 680 685  
 Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro  
 690 695 700

Glu	Leu	Tyr	Phe	Val	Lys	Val	Asp	Val	Thr	Gly	Ala	Tyr	Asp	Thr	Ile
705					710					715					720
Pro	Gln	Asp	Arg	Leu	Thr	Glu	Val	Ile	Ala	Ser	Ile	Ile	Lys	Pro	Gln
				725					730					735	
Asn	Thr	Tyr	Cys	Val	Arg	Arg	Tyr	Ala	Val	Val	Gln	Lys	Ala	Ala	His
			740					745					750		
Gly	His	Val	Arg	Lys	Ala	Phe	Lys	Ser	His	Val	Ser	Thr	Leu	Thr	Asp
		755					760					765			
Leu	Gln	Pro	Tyr	Met	Arg	Gln	Phe	Val	Ala	His	Leu	Gln	Glu	Thr	Ser
	770					775					780				
Pro	Leu	Arg	Asp	Ala	Val	Val	Ile	Glu	Gln	Ser	Ser	Ser	Leu	Asn	Glu
785					790					795					800
Ala	Ser	Ser	Gly	Leu	Phe	Asp	Val	Phe	Leu	Arg	Phe	Met	Cys	His	His
				805					810					815	
Ala	Val	Arg	Ile	Arg	Gly	Lys	Ser	Tyr	Val	Gln	Cys	Gln	Gly	Ile	Pro
			820					825					830		
Gln	Gly	Ser	Ile	Leu	Ser	Thr	Leu	Leu	Cys	Ser	Leu	Cys	Tyr	Gly	Asp
	835						840					845			
Met	Glu	Asn	Lys	Leu	Phe	Ala	Gly	Ile	Arg	Arg	Asp	Gly	Leu	Leu	Leu
	850					855					860				
Arg	Leu	Val	Asp	Asp	Phe	Leu	Leu	Val	Thr	Pro	His	Leu	Thr	His	Ala
865					870					875					880
Lys	Thr	Phe	Leu	Arg	Thr	Leu	Val	Arg	Gly	Val	Pro	Glu	Tyr	Gly	Cys
			885						890					895	
Val	Val	Asn	Leu	Arg	Lys	Thr	Val	Val	Asn	Phe	Pro	Val	Glu	Asp	Glu
			900					905					910		
Ala	Leu	Gly	Gly	Thr	Ala	Phe	Val	Gln	Met	Pro	Ala	His	Gly	Leu	Phe
		915					920					925			
Pro	Trp	Cys	Gly	Leu	Leu	Leu	Asp	Thr	Arg	Thr	Leu	Glu	Val	Gln	Ser
	930					935					940				
Asp	Tyr	Ser	Ser	Tyr	Ala	Arg	Thr	Ser	Ile	Arg	Ala	Ser	Leu	Thr	Phe
945					950					955					960
Asn	Arg	Gly	Phe	Lys	Ala	Gly	Arg	Asn	Met	Arg	Arg	Lys	Leu	Phe	Gly
				965					970					975	
Val	Leu	Arg	Leu	Lys	Cys	His	Ser	Leu	Phe	Leu	Asp	Leu	Gln	Val	Asn
			980					985					990		
Ser	Leu	Gln	Thr	Val	Cys	Thr	Asn	Ile	Tyr	Lys	Ile	Leu	Leu	Leu	Gln
		995					1000					1005			
Ala	Tyr	Arg	Phe	His	Ala	Cys	Val	Leu	Gln	Leu	Pro	Phe	His	Gln	Gln
	1010					1015					1020				
Val	Trp	Lys	Asn	Pro	Thr	Phe	Phe	Leu	Arg	Val	Ile	Ser	Asp	Thr	Ala
1025					1030					1035					1040

Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser Leu  
 1045 1050 1055  
 Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln Trp  
 1060 1065 1070  
 Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg Val Thr  
 1075 1080 1085  
 Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr Gln Leu Ser  
 1090 1095 1100  
 Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala Ala Ala Asn  
 1105 1110 1115 1120  
 Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp Leu Glu Gln Lys  
 1125 1130 1135  
 Leu Ile Ser Glu Glu Asp Leu Asn Ser Ala Val Asp His His His His  
 1140 1145 1150  
 His His

(2) INFORMATION FOR SEQ ID NO:324:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1200 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:324:

Met Pro Arg Gly Ser His His His His His His Gly Met Ala Ser Met  
 1 5 10 15  
 Thr Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Leu  
 20 25 30  
 Asp Pro Ser Ser Arg Ser Ala Ala Gly Thr Met Glu Phe Ala Ala Ala  
 35 40 45  
 Ser Thr Gln Arg Cys Val Leu Leu Arg Thr Trp Glu Ala Leu Ala Pro  
 50 55 60  
 Ala Thr Pro Ala Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser  
 65 70 75 80  
 Leu Leu Arg Ser His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val  
 85 90 95  
 Arg Arg Leu Gly Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro  
 100 105 110  
 Ala Ala Phe Arg Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp  
 115 120 125  
 Asp Ala Arg Pro Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys

130					135					140					
Leu 145	Lys	Glu	Leu	Val	Ala 150	Arg	Val	Leu	Gln	Arg 155	Leu	Cys	Glu	Arg	Gly 160
Ala	Lys	Asn	Val	Leu 165	Ala	Phe	Gly	Phe	Ala 170	Leu	Leu	Asp	Gly	Ala	Arg 175
Gly	Gly	Pro	Pro 180	Glu	Ala	Phe	Thr	Thr 185	Ser	Val	Arg	Ser	Tyr	Leu	Pro 190
Asn	Thr	Val 195	Thr	Asp	Ala	Leu	Arg 200	Gly	Ser	Gly	Ala	Trp	Gly	Leu	Leu 205
Leu	Arg 210	Arg	Val	Gly	Asp	Asp 215	Val	Leu	Val	His	Leu	Leu	Ala	Arg	Cys 220
Ala 225	Leu	Phe	Val	Leu	Val 230	Ala	Pro	Ser	Cys	Ala 235	Tyr	Gln	Val	Cys	Gly 240
Pro	Pro	Leu	Tyr	Gln 245	Leu	Gly	Ala	Ala	Thr 250	Gln	Ala	Arg	Pro	Pro	Pro 255
His	Ala	Ser	Gly 260	Pro	Arg	Arg	Arg	Leu 265	Gly	Cys	Glu	Arg	Ala	Trp	Asn 270
His	Ser	Val 275	Arg	Glu	Ala	Gly	Val 280	Pro	Leu	Gly	Leu	Pro	Ala	Pro	Gly 285
Ala	Arg 290	Arg	Arg	Gly	Gly	Ser 295	Ala	Ser	Arg	Ser	Leu	Pro	Leu	Pro	Lys 300
Arg 305	Pro	Arg	Arg	Gly	Ala 310	Ala	Pro	Glu	Pro	Glu 315	Arg	Thr	Pro	Val	Gly 320
Gln	Gly	Ser	Trp	Ala 325	His	Pro	Gly	Arg	Thr 330	Arg	Gly	Pro	Ser	Asp	Arg 335
Gly	Phe	Cys	Val 340	Val	Ser	Pro	Ala	Arg 345	Pro	Ala	Glu	Glu	Ala	Thr	Ser 350
Leu	Glu	Gly 355	Ala	Leu	Ser	Gly	Thr 360	Arg	His	Ser	His	Pro	Ser	Val	Gly 365
Arg	Gln 370	His	His	Ala	Gly	Pro 375	Pro	Ser	Thr	Ser	Arg 380	Pro	Pro	Arg	Pro 385
Trp 385	Asp	Thr	Pro	Cys	Pro 390	Pro	Val	Tyr	Ala	Glu 395	Thr	Lys	His	Phe	Leu 400
Tyr	Ser	Ser	Gly	Asp 405	Lys	Glu	Gln	Leu	Arg 410	Pro	Ser	Phe	Leu	Leu	Ser 415
Ser	Leu	Arg	Pro 420	Ser	Leu	Thr	Gly	Ala 425	Arg	Arg	Leu	Val	Glu	Thr	Ile 430
Phe	Leu	Gly 435	Ser	Arg	Pro	Trp	Met 440	Pro	Gly	Thr	Pro	Arg 445	Arg	Leu	Pro 450
Arg	Leu 450	Pro	Gln	Arg	Tyr	Trp 455	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu 460

Leu Gly Asn His Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His  
 465 470 475 480  
 Cys Pro Leu Arg Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg  
 485 490 495  
 Glu Lys Pro Gln Gly Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp  
 500 505 510  
 Pro Arg Arg Leu Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln  
 515 520 525  
 Val Tyr Gly Phe Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly  
 530 535 540  
 Leu Trp Gly Ser Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys  
 545 550 555 560  
 Lys Phe Ile Ser Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu  
 565 570 575  
 Thr Trp Lys Met Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro  
 580 585 590  
 Gly Val Gly Cys Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile  
 595 600 605  
 Leu Ala Lys Phe Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu  
 610 615 620  
 Leu Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg  
 625 630 635 640  
 Leu Phe Phe Tyr Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly  
 645 650 655  
 Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala  
 660 665 670  
 Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg  
 675 680 685  
 Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met  
 690 695 700  
 Asp Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu  
 705 710 715 720  
 Arg Leu Thr Ser Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu  
 725 730 735  
 Arg Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp  
 740 745 750  
 Asp Ile His Arg Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln  
 755 760 765  
 Asp Pro Pro Pro Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala  
 770 775 780  
 Tyr Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile  
 785 790 795 800



Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln  
 805 810 815  
 Lys Ala Ala His Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser  
 820 825 830  
 Thr Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu  
 835 840 845  
 Gln Glu Thr Ser Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser  
 850 855 860  
 Ser Leu Asn Glu Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe  
 865 870 875 880  
 Met Cys His His Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys  
 885 890 895  
 Gln Gly Ile Pro Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu  
 900 905 910  
 Cys Tyr Gly Asp Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp  
 915 920 925  
 Gly Leu Leu Leu Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His  
 930 935 940  
 Leu Thr His Ala Lys Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro  
 945 950 955 960  
 Glu Tyr Gly Cys Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro  
 965 970 975  
 Val Glu Asp Glu Ala Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala  
 980 985 990  
 His Gly Leu Phe Pro Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu  
 995 1000 1005  
 Glu Val Gln Ser Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala  
 1010 1015 1020  
 Ser Leu Thr Phe Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg  
 1025 1030 1035 1040  
 Lys Leu Phe Gly Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp  
 1045 1050 1055  
 Leu Gln Val Asn Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile  
 1060 1065 1070  
 Leu Leu Leu Gln Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro  
 1075 1080 1085  
 Phe His Gln Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile  
 1090 1095 1100  
 Ser Asp Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala  
 1105 1110 1115 1120  
 Gly Met Ser Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu  
 1125 1130 1135

Ala Val Gln Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg  
1140 1145 1150  
His Arg Val Thr Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln  
1155 1160 1165  
Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu  
1170 1175 1180  
Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp  
1185 1190 1195 1200

(2) INFORMATION FOR SEQ ID NO:325:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1189 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:325:

Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile  
1 5 10 15  
Ser Tyr Ile Tyr Ala Asp Pro Ser Ser Arg Ser Ala Ala Gly Thr Met  
20 25 30  
Glu Phe Ala Ala Ala Ser Thr Gln Arg Cys Val Leu Leu Arg Thr Trp  
35 40 45  
Glu Ala Leu Ala Pro Ala Thr Pro Ala Met Pro Arg Ala Pro Arg Cys  
50 55 60  
Arg Ala Val Arg Ser Leu Arg Ser His Tyr Arg Glu Val Leu Pro  
65 70 75 80  
Leu Ala Thr Phe Val Arg Arg Leu Gly Pro Gln Gly Trp Arg Leu Val  
85 90 95  
Gln Arg Gly Asp Pro Ala Ala Phe Arg Ala Leu Val Ala Gln Cys Leu  
100 105 110  
Val Cys Val Pro Trp Asp Ala Arg Pro Pro Pro Ala Ala Pro Ser Phe  
115 120 125  
Arg Gln Val Ser Cys Leu Lys Glu Leu Val Ala Arg Val Leu Gln Arg  
130 135 140  
Leu Cys Glu Arg Gly Ala Lys Asn Val Leu Ala Phe Gly Phe Ala Leu  
145 150 155 160  
Leu Asp Gly Ala Arg Gly Gly Pro Pro Glu Ala Phe Thr Thr Ser Val  
165 170 175  
Arg Ser Tyr Leu Pro Asn Thr Val Thr Asp Ala Leu Arg Gly Ser Gly  
180 185 190

Ala Trp Gly Leu Leu Leu Arg Arg Val Gly Asp Asp Val Leu Val His  
 195 200 205  
 Leu Leu Ala Arg Cys Ala Leu Phe Val Leu Val Ala Pro Ser Cys Ala  
 210 215 220  
 Tyr Gln Val Cys Gly Pro Pro Leu Tyr Gln Leu Gly Ala Ala Thr Gln  
 225 230 235 240  
 Ala Arg Pro Pro Pro His Ala Ser Gly Pro Arg Arg Arg Leu Gly Cys  
 245 250 255  
 Glu Arg Ala Trp Asn His Ser Val Arg Glu Ala Gly Val Pro Leu Gly  
 260 265 270  
 Leu Pro Ala Pro Gly Ala Arg Arg Gly Gly Ser Ala Ser Arg Ser  
 275 280 285  
 Leu Pro Leu Pro Lys Arg Pro Arg Arg Gly Ala Ala Pro Glu Pro Glu  
 290 295 300  
 Arg Thr Pro Val Gly Gln Gly Ser Trp Ala His Pro Gly Arg Thr Arg  
 305 310 315 320  
 Gly Pro Ser Asp Arg Gly Phe Cys Val Val Ser Pro Ala Arg Pro Ala  
 325 330 335  
 Glu Glu Ala Thr Ser Leu Glu Gly Ala Leu Ser Gly Thr Arg His Ser  
 340 345 350  
 His Pro Ser Val Gly Arg Gln His His Ala Gly Pro Pro Ser Thr Ser  
 355 360 365  
 Arg Pro Pro Arg Pro Trp Asp Thr Pro Cys Pro Pro Val Tyr Ala Glu  
 370 375 380  
 Thr Lys His Phe Leu Tyr Ser Ser Gly Asp Lys Glu Gln Leu Arg Pro  
 385 390 395 400  
 Ser Phe Leu Leu Ser Ser Leu Arg Pro Ser Leu Thr Gly Ala Arg Arg  
 405 410 415  
 Leu Val Glu Thr Ile Phe Leu Gly Ser Arg Pro Trp Met Pro Gly Thr  
 420 425 430  
 Pro Arg Arg Leu Pro Arg Leu Pro Gln Arg Tyr Trp Gln Met Arg Pro  
 435 440 445  
 Leu Phe Leu Glu Leu Leu Gly Asn His Ala Gln Cys Pro Tyr Gly Val  
 450 455 460  
 Leu Leu Lys Thr His Cys Pro Leu Arg Ala Ala Val Thr Pro Ala Ala  
 465 470 475 480  
 Gly Val Cys Ala Arg Glu Lys Pro Gln Gly Ser Val Ala Ala Pro Glu  
 485 490 495  
 Glu Glu Asp Thr Asp Pro Arg Arg Leu Val Gln Leu Leu Arg Gln His  
 500 505 510  
 Ser Ser Pro Trp Gln Val Tyr Gly Phe Val Arg Ala Cys Leu Arg Arg  
 515 520 525

Leu Val Pro Pro Gly Leu Trp Gly Ser Arg His Asn Glu Arg Arg Phe  
 530 535 540  
 Leu Arg Asn Thr Lys Lys Phe Ile Ser Leu Gly Lys His Ala Lys Leu  
 545 550 555 560  
 Ser Leu Gln Glu Leu Thr Trp Lys Met Ser Val Arg Asp Cys Ala Trp  
 565 570 575  
 Leu Arg Arg Ser Pro Gly Val Gly Cys Val Pro Ala Ala Glu His Arg  
 580 585 590  
 Leu Arg Glu Glu Ile Leu Ala Lys Phe Leu His Trp Leu Met Ser Val  
 595 600 605  
 Tyr Val Val Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr  
 610 615 620  
 Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val Trp Ser Lys  
 625 630 635 640  
 Leu Gln Ser Ile Gly Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg  
 645 650 655  
 Glu Leu Ser Glu Ala Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala  
 660 665 670  
 Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg  
 675 680 685  
 Pro Ile Val Asn Met Asp Tyr Val Val Gly Ala Arg Thr Phe Arg Arg  
 690 695 700  
 Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg Val Lys Ala Leu Phe Ser  
 705 710 715 720  
 Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser  
 725 730 735  
 Val Leu Gly Leu Asp Asp Ile His Arg Ala Trp Arg Thr Phe Val Leu  
 740 745 750  
 Arg Val Arg Ala Gln Asp Pro Pro Pro Glu Leu Tyr Phe Val Lys Val  
 755 760 765  
 Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu  
 770 775 780  
 Val Ile Ala Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg  
 785 790 795 800  
 Tyr Ala Val Val Gln Lys Ala Ala His Gly His Val Arg Lys Ala Phe  
 805 810 815  
 Lys Ser His Val Ser Thr Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln  
 820 825 830  
 Phe Val Ala His Leu Gln Glu Thr Ser Pro Leu Arg Asp Ala Val Val  
 835 840 845  
 Ile Glu Gln Ser Ser Ser Leu Asn Glu Ala Ser Ser Gly Leu Phe Asp  
 850 855 860

Val Phe Leu Arg Phe Met Cys His His Ala Val Arg Ile Arg Gly Lys  
 865 870 875 880  
 Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln Gly Ser Ile Leu Ser Thr  
 885 890 895  
 Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met Glu Asn Lys Leu Phe Ala  
 900 905 910  
 Gly Ile Arg Arg Asp Gly Leu Leu Leu Arg Leu Val Asp Asp Phe Leu  
 915 920 925  
 Leu Val Thr Pro His Leu Thr His Ala Lys Thr Phe Leu Arg Thr Leu  
 930 935 940  
 Val Arg Gly Val Pro Glu Tyr Gly Cys Val Val Asn Leu Arg Lys Thr  
 945 950 955 960  
 Val Val Asn Phe Pro Val Glu Asp Glu Ala Leu Gly Gly Thr Ala Phe  
 965 970 975  
 Val Gln Met Pro Ala His Gly Leu Phe Pro Trp Cys Gly Leu Leu Leu  
 980 985 990  
 Asp Thr Arg Thr Leu Glu Val Gln Ser Asp Tyr Ser Ser Tyr Ala Arg  
 995 1000 1005  
 Thr Ser Ile Arg Ala Ser Leu Thr Phe Asn Arg Gly Phe Lys Ala Gly  
 1010 1015 1020  
 Arg Asn Met Arg Arg Lys Leu Phe Gly Val Leu Arg Leu Lys Cys His  
 1025 1030 1035 1040  
 Ser Leu Phe Leu Asp Leu Gln Val Asn Ser Leu Gln Thr Val Cys Thr  
 1045 1050 1055  
 Asn Ile Tyr Lys Ile Leu Leu Leu Gln Ala Tyr Arg Phe His Ala Cys  
 1060 1065 1070  
 Val Leu Gln Leu Pro Phe His Gln Gln Val Trp Lys Asn Pro Thr Phe  
 1075 1080 1085  
 Phe Leu Arg Val Ile Ser Asp Thr Ala Ser Leu Cys Tyr Ser Ile Leu  
 1090 1095 1100  
 Lys Ala Lys Asn Ala Gly Met Ser Leu Gly Ala Lys Gly Ala Ala Gly  
 1105 1110 1115 1120  
 Pro Leu Pro Ser Glu Ala Val Gln Trp Leu Cys His Gln Ala Phe Leu  
 1125 1130 1135  
 Leu Lys Leu Thr Arg His Arg Val Thr Tyr Val Pro Leu Leu Gly Ser  
 1140 1145 1150  
 Leu Arg Thr Ala Gln Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr  
 1155 1160 1165  
 Leu Thr Ala Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe  
 1170 1175 1180  
 Lys Thr Ile Leu Asp  
 1185

(2) INFORMATION FOR SEQ ID NO:326:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 58 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:326:

TGCGCACGTG GGAAGCCCTG GCAGATCTGA ATTCCACCAT GCCGCGCGCT CCCCCTG 58

(2) INFORMATION FOR SEQ ID NO:327:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 92 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:327:

CTGCCCTCAG ACTTCAAGAC CATCCTGGAC TACAAGGACG ACGATGACAA ATGAATTCAG 60  
ATCTGCGGCC GCCACCGCGG TGGAGCTCCA GC 92

(2) INFORMATION FOR SEQ ID NO:328:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:328:

CGGGACGGGC TGCTCCTGCG TTTGGTGGAC GCGTTCTTGT TGGTGACACC TCACCTCACC 60

(2) INFORMATION FOR SEQ ID NO:329:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 62 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:329:

ATTCCGTCGA GCAGAGTTAG GGTTAGGGTT AGGGTTAGGG TTAGGGTTAG GGTTAGGGTT 60  
AG 62

(2) INFORMATION FOR SEQ ID NO:330:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 84 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:330:

GGGAGATCTT AATACGACTC ACTATAGATT CAGGCCATGG TGCTGCGCCG GCTGTCAGGC 60  
TCCCACGACG TAGTCCATGT TCAC 84

(2) INFORMATION FOR SEQ ID NO:331:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:331:

GGGTCTAGAT CCGGAAGAGT GTCTGGAGCA AG 32

(2) INFORMATION FOR SEQ ID NO:332:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 84 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:332:

GGGAGATCTT AATACGACTC ACTATAGATT CAGGCCATGG TGCTGCGCCG GCTGTCAGGG 60  
CGGCCTTCTG GACCACGGCA TACC 84

(2) INFORMATION FOR SEQ ID NO:333:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:333:

GGTCTAGACG ATATCCACAG GGCCTGGCGC

30

(2) INFORMATION FOR SEQ ID NO:334:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1407 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:334:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
1				5					10					15		
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
			35				40					45				
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65					70				75					80		
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
			85					90						95		
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
			115				120					125				
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
	130					135					140					
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
145					150					155					160	
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165				170						175		
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180					185					190			
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
			195				200						205			



Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe  
 210 215 220  
 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser  
 225 230 235 240  
 Gly Arg Thr Gln Ile Ser Ser Ser Ser Phe Glu Phe Ala Ala Ala Ser  
 245 250 255  
 Thr Gln Arg Cys Val Leu Leu Arg Thr Trp Glu Ala Leu Ala Pro Ala  
 260 265 270  
 Thr Pro Ala Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu  
 275 280 285  
 Leu Arg Ser His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg  
 290 295 300  
 Arg Leu Gly Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala  
 305 310 315 320  
 Ala Phe Arg Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp  
 325 330 335  
 Ala Arg Pro Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu  
 340 345 350  
 Lys Glu Leu Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala  
 355 360 365  
 Lys Asn Val Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly  
 370 375 380  
 Gly Pro Pro Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn  
 385 390 395 400  
 Thr Val Thr Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu  
 405 410 415  
 Arg Arg Val Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala  
 420 425 430  
 Leu Phe Val Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro  
 435 440 445  
 Pro Leu Tyr Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His  
 450 455 460  
 Ala Ser Gly Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His  
 465 470 475 480  
 Ser Val Arg Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala  
 485 490 495  
 Arg Arg Arg Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg  
 500 505 510  
 Pro Arg Arg Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln  
 515 520 525  
 Gly Ser Trp Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly  
 530 535 540

Phe	Cys	Val	Val	Ser	Pro	Ala	Arg	Pro	Ala	Glu	Glu	Ala	Thr	Ser	Leu
545					550					555					560
Glu	Gly	Ala	Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg
				565					570					575	
Gln	His	His	Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp
			580					585					590		
Asp	Thr	Pro	Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr
		595					600					605			
Ser	Ser	Gly	Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser
	610					615					620				
Leu	Arg	Pro	Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe
625					630					635					640
Leu	Gly	Ser	Arg	Pro	Trp	Met	Pro	Gly	Thr	Pro	Arg	Arg	Leu	Pro	Arg
				645					650					655	
Leu	Pro	Gln	Arg	Tyr	Trp	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu	Leu
			660					665					670		
Gly	Asn	His	Ala	Gln	Cys	Pro	Tyr	Gly	Val	Leu	Leu	Lys	Thr	His	Cys
		675					680					685			
Pro	Leu	Arg	Ala	Ala	Val	Thr	Pro	Ala	Ala	Gly	Val	Cys	Ala	Arg	Glu
	690					695					700				
Lys	Pro	Gln	Gly	Ser	Val	Ala	Ala	Pro	Glu	Glu	Glu	Asp	Thr	Asp	Pro
705					710					715					720
Arg	Arg	Leu	Val	Gln	Leu	Leu	Arg	Gln	His	Ser	Ser	Pro	Trp	Gln	Val
				725					730					735	
Tyr	Gly	Phe	Val	Arg	Ala	Cys	Leu	Arg	Arg	Leu	Val	Pro	Pro	Gly	Leu
			740					745					750		
Trp	Gly	Ser	Arg	His	Asn	Glu	Arg	Arg	Phe	Leu	Arg	Asn	Thr	Lys	Lys
		755					760					765			
Phe	Ile	Ser	Leu	Gly	Lys	His	Ala	Lys	Leu	Ser	Leu	Gln	Glu	Leu	Thr
	770					775					780				
Trp	Lys	Met	Ser	Val	Arg	Asp	Cys	Ala	Trp	Leu	Arg	Arg	Ser	Pro	Gly
785					790					795					800
Val	Gly	Cys	Val	Pro	Ala	Ala	Glu	His	Arg	Leu	Arg	Glu	Glu	Ile	Leu
				805					810					815	
Ala	Lys	Phe	Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu
			820					825					830		
Arg	Ser	Phe	Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu
		835					840					845			
Phe	Phe	Tyr	Arg	Pro	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile
	850					855				860					
Arg	Gln	His	Leu	Lys	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser	Glu	Ala	Glu
865					870					875					880

Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu  
 885 890 895  
 Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp  
 900 905 910  
 Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg  
 915 920 925  
 Leu Thr Ser Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg  
 930 935 940  
 Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp  
 945 950 955 960  
 Ile His Arg Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp  
 965 970 975  
 Pro Pro Pro Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr  
 980 985 990  
 Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile  
 995 1000 1005  
 Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys  
 1010 1015 1020  
 Ala Ala His Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr  
 1025 1030 1035 1040  
 Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln  
 1045 1050 1055  
 Glu Thr Ser Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser  
 1060 1065 1070  
 Leu Asn Glu Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met  
 1075 1080 1085  
 Cys His His Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln  
 1090 1095 1100  
 Gly Ile Pro Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys  
 1105 1110 1115 1120  
 Tyr Gly Asp Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly  
 1125 1130 1135  
 Leu Leu Leu Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu  
 1140 1145 1150  
 Thr His Ala Lys Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu  
 1155 1160 1165  
 Tyr Gly Cys Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val  
 1170 1175 1180  
 Glu Asp Glu Ala Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His  
 1185 1190 1195 1200  
 Gly Leu Phe Pro Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu  
 1205 1210 1215

Val Gln Ser Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser  
 1220 1225 1230  
 Val Thr Phe Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys  
 1235 1240 1245  
 Leu Phe Gly Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu  
 1250 1255 1260  
 Gln Val Asn Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu  
 1265 1270 1275 1280  
 Leu Leu Gln Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe  
 1285 1290 1295  
 His Gln Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser  
 1300 1305 1310  
 Asp Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly  
 1315 1320 1325  
 Met Ser Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala  
 1330 1335 1340  
 Val Gln Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His  
 1345 1350 1355 1360  
 Arg Val Thr Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr  
 1365 1370 1375  
 Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala  
 1380 1385 1390  
 Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp  
 1395 1400 1405

(2) INFORMATION FOR SEQ ID NO:335:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:335:

Gly Ser Thr His Ile Ser His Ile Ser His Ile Ser His Ile Ser His  
 1 5 10 15  
 Ile Ser His Ile Ser His Ile Ser His Ile Ser  
 20 25